The Impact on Emotion Classification Performance and Gaze Behavior of Foveal vs. Extrafoveal Processing of Facial Features

Anthony P. Atkinson

Durham University, UK

Hannah E. Smithson

University of Oxford, UK

Author Note

The authors are grateful to İlayda Türk for running a pilot study, to Sophie Apps, Ayşe Konuk, Dima Lin, Helena Mortimer, and Hannah White for assistance with data collection, to Elaine Stanton for initial help in processing the eye-tracking data, and to the two reviewers and editor for their helpful feedback. Data and associated code can be downloaded from https://osf.io/vduzt/

Correspondence concerning this article should be addressed to Anthony P. Atkinson, Department of Psychology, Durham University, Science Laboratories, South Road, Durham, DH1 3LE, UK. Email: a.p.atkinson@durham.ac.uk

Abstract

At normal interpersonal distances all features of a face cannot fall within one's fovea simultaneously. Given that certain facial features are differentially informative of different emotions, does the ability to identify facially expressed emotions vary according to the feature fixated and do saccades preferentially seek diagnostic features? Previous findings are equivocal. We presented faces for a brief time, insufficient for a saccade, at a spatial position that guaranteed that a given feature – an eye, cheek, the central brow, or mouth – fell at the fovea. Across two experiments, observers were more accurate and faster at discriminating angry expressions when the high spatial-frequency information of the brow was projected to their fovea than when one or other cheek or eye was. Performance in classifying fear and happiness (Experiment 1) was not influenced by whether the most informative features (eyes and mouth, respectively) were projected foveally or extrafoveally. Observers more accurately distinguished between fearful and surprised expressions (Experiment 2) when the mouth was projected to the fovea. Reflexive first saccades tended towards the left and center of the face rather than preferentially targeting emotion-distinguishing features. These results reflect the integration of task-relevant information across the face constrained by the differences between foveal and extrafoveal processing (Peterson & Eckstein, 2012).

Keywords: facial expression; emotion recognition; fixation; eye movements; peripheral vision

Public significance statement

Different parts of the face provide important cues about underlying emotional states (e.g., the furrowed brow of an angry face, the shape of the mouth in fear vs. surprise). This study shows that a single fixation on such a diagnostic feature can enhance the ability to recognize the emotion, relative to fixating another part of the face; yet, when diagnostic features are in the visual periphery, one's eyes do not automatically seek them out.

The Impact on Emotion Classification Performance and Gaze Behavior of Foveal vs. Extrafoveal Processing of Facial Features

This study focuses on the differential contributions of foveal and extrafoveal processing of facial features to the identification of facially expressed emotions. The motivation for this study comes from the confluence of three facts, which we elaborate in more detail below: (a) There are both quantitative and qualitative differences between foveal and extrafoveal visual processing, (b) at normal interpersonal distances not all features of one person's face can fall within another person's fovea at once, and (c) certain facial features carry information 'diagnostic' of specific emotions. We tested two main hypotheses that follow from these facts: (1) Fixation on an emotion-distinguishing facial feature for which medium-to-high spatial frequency information is most informative will enhance emotion identification performance compared to when another part of the face is fixated causing the diagnostic feature to appear in extrafoveal vision. (2) When required to identify the expressed emotion, observers' initial eye movements will reflect the location of task-relevant features – specifically, they will preferentially saccade toward emotion-distinguishing facial features, especially those features for which medium-to-high spatial frequency information would be most informative.

The fovea, a small region of the retina that corresponds to the central 1.7° of the visual field (Wandell, 1995)[1], is preferentially specialized for processing fine spatial detail. With increasing eccentricity from the fovea, there is a decline in both visual acuity (i.e., the spatial resolving capacity of the visual system) and contrast sensitivity (i.e., the ability to detect differences in contrast) (Robson & Graham, 1981; Rosenholtz, 2016). Peripheral vision also differs qualitatively from central vision, receiving different processing and optimized for different tasks (Rosenholtz, 2016; Strasburger, Rentschler, & Jüttner, 2011).

The average height of an adult human face is approximately 18cm (Fang, Clapham, & Chung, 2011). At what Hall (1966) calls the "close phase of personal distance" (~76-45 cm), a face will thus subtend visual angles of 13.4-22°, and at far personal distances (~122-76 cm), 8.4-13.4°. Therefore, during many everyday face-to-face interactions, a fixation on someone's eye, for example, will mean that much of the rest of that person's face will fall outside of the viewer's fovea. Under such conditions, detailed vision of another's face thus requires fixations on multiple features, which fall mostly on the eyes, nose, and mouth (Arizpe, Walsh, Yovel, & Baker, 2017; Bindemann, Scheepers, & Burton, 2009; Henderson, Williams, & Falk, 2005; Hsiao & Cottrell, 2008; Jack, Blais, Scheepers, Schyns, & Caldara, 2009; Mehoudar, Arizpe, Baker, & Yovel, 2014; Peterson & Eckstein, 2013; Yarbus, 1967). Features falling outside the fovea nevertheless receive some visual processing, perhaps determining the next fixation location or even contributing directly to the extraction of socially relevant information, such as identity and emotion. In the present study, we focus on the relative contributions to facial emotion perception of foveated and non-foveated facial features.

Findings from studies that involved presenting observers with face images filtered with randomly located Gaussian apertures or "Bubbles", rendering only parts of the face clearly visible, have shown that specific facial features carry information 'diagnostic' of specific emotions[2]; for example, fear and surprise are revealed by wide-open eyes, anger by a furrowed brow; the mouth is diagnostic of happiness and differentiates fear from surprise (Smith, Cottrell, Gosselin, & Schyns, 2005; Smith & Merlusca, 2014). These findings confirmed and extended earlier research showing that the ability to recognize particular facially expressed emotions varies depending on which features are visible when participants are presented with partial faces or isolated face parts (e.g., Boucher & Ekman, 1975; Calder, Young, Keane, & Dean, 2000). Even so, the spatial relations between features and the context of the whole face are also important for perceiving its emotional expression (e.g., Calder, Young, et al., 2000; Calvo & Beltrán, 2014; Tanaka, Kaiser, Butler, & Le Grand, 2012; White, 2000). Indeed, there is some evidence suggesting that the relative contribution of configural and holistic processing to face perception is greater at normal interpersonal distances than at larger distances (Oruç & Barton, 2010; Ross & Gauthier, 2015; Yang, Shafai, & Oruc, 2014) (though see McKone, 2009). The present study focuses on the processing of specific features in the context of a whole, unaltered face, where the only filtering is that done by the eye.

Experimental separation of foveal and extrafoveal contributions to face processing

Studies of face and emotion perception typically employ free-viewing conditions in which the observers can make one or more fixations on the image. Free-viewing conditions do not readily allow the teasing apart of foveal and extrafoveal visual processing so that their contributions to task performance can be examined separately. Gaze-contingent windowing to reveal parts of faces can be used to offer stimuli to only foveal or extrafoveal retina, but these methods necessarily disrupt holistic face processing (particularly windowing that offers stimuli to only extrafoveal retina; e.g., Van Belle, De Graef, Verfaillie, Rossion, & Lefèvre, 2010). A simple way of controlling presentation to foveal and extrafoveal visual fields without disrupting holistic processing – a method that we use here – is to present stimuli only briefly: Since a finite time is required to program and initiate a saccade, presentation and removal of the image can be completed before an eye-movement can redirect the fovea to a new location on the image. This manipulation has been used in a number of studies of facial emotion perception (Boll & Gamer, 2014; Gamer & Büchel, 2009; Gamer, Schmitz, Tittgemeyer, & Schilbach, 2013; Gamer, Zurowski, & Büchel, 2010; Kliemann, Dziobek, Hatri, Baudewig, & Heekeren, 2012; Kliemann, Dziobek, Hatri, Steimke, & Heekeren, 2010; Neath & Itier, 2014; Scheller, Büchel, & Gamer, 2012). With the exception of one these studies (Neath & Itier, 2014), all have used stimulus presentation times of 150 or 200 ms (Neath & Itier used presentation times of 16.67, 50 and 100 ms). However, the required brevity of stimulus presentation is contentious. Regular saccade latencies are of the order of 135-220ms, but this includes time for fixation neurons of the superior colliculus to disengage (Fuchs, Kaneko, & Scudder, 1985; Wurtz, 1996) and removal of the fixation stimulus prior to presentation of the target, as was done in all of the cited studies using emotional faces as stimuli, may shorten the critical window for stimulus presentation to 90-120 ms (Saslow, 1967; Walker, Kentridge, & Findlay, 1995), though not for some tasks, including reading (Liversedge et al., 2004). To be safe, in the present study we enforced fixation on specific facial locations by presenting faces in a fixation-contingent manner for approximately 80 ms.

Does emotion classification performance vary as a function of initial fixation?

In emotion discrimination or classification tasks, how much, if any, advantage is provided by fixation specifically on emotion-distinguishing features? There are certainly some cases in which successful emotion recognition depends on the fixation of certain facial regions. Notably, a selective impairment in recognizing fear from faces associated with bilateral amygdala damage is the result of a failure to saccade spontaneously to and thus fixate the eye region (Adolphs et al., 2005), a region that is informative for fear (Smith, et al., 2005; Smith & Merlusca, 2014). Remarkably, instructing the patient with bilateral amygdala damage to fixate the eyes restored fear recognition performance to normal levels (Adolphs, et al., 2005). There is some evidence that, when required to judge the emotional content of facial expressions under free-viewing conditions, neurologically healthy observers tend to spend more time fixating different regions of the face depending on the viewed emotion (e.g., Beaudry, Roy-Charland, Perron, Cormier, & Tapp, 2013; Schurgin et al., 2014) and that accuracy in detecting emotional expressions is predicted by participants' fixation patterns, though mostly for subtle rather than strong expressions (Vaidya, Jin, & Fellows, 2014). Nevertheless, the facial regions that are fixated more often or for longer in these studies (Beaudry, et al., 2013; Schurgin, et al., 2014; Vaidya, et al., 2014) do not always line up neatly with those features identified as emotion-distinguishing (Smith, et al., 2005; Smith & Merlusca, 2014; Vaidya, et al., 2014).

Given that the first one or two fixations are most critical for the discrimination of facial emotional expression (Schurgin, et al., 2014) and identity (Hsiao & Cottrell, 2008), is a single fixation on an emotion-distinguishing facial feature sufficient to enhance emotion identification performance compared to when that feature appears in extrafoveal vision? The evidence is mixed. Using the brief-fixation technique, Gamer and Büchel (2009) had participants view fearful, angry, happy and emotionally neutral faces with fixation enforced at either the center of the mouth or at one or other of the eyes. The faces subtended an average of 13.6° vertically. The participants were marginally more accurate in classifying fearful and angry expressions when fixating an eye than when fixating the mouth (though these effects were not statistically significant) and were equally accurate in classifying happy faces when fixating an eye or the mouth. In four experiments across three subsequent studies by the same research group (Boll & Gamer, 2014; Gamer, et al., 2010; Scheller, et al., 2012) again there were no statistically significant effects of initial fixation on emotion identification accuracy. Gamer et al. (2010) did, though, find that participants were faster to classify happiness with fixation on the center of the mouth and to classify fear with fixation on the midpoint between the eyes. Note that, in these latter three studies, the eyes themselves were not fixation locations, but instead, a point between the eyes, directly above the mouth. Neath and Itier (2014) also found no effect on emotion classification accuracy of initial fixation location. Their experiment used disgusted, surprised, fearful, happy and neutral faces and fixation locations comprising the chin, nose, center of forehead and cheeks, as well as the eyes and mouth center. The faces were of a size corresponding to close personal distance (images subtended an average of 26.8° vertically).[3]

Two other studies using the brief-fixation paradigm did find significant effects of initial fixation location on emotion classification performance, however. Viewing images in which fearful, happy and neutral faces subtended a visual angle of approximately 14° vertically**,** Kliemann et al.'s (2010) participants were more accurate and faster to classify fearful faces with fixation on the central point between the eyes compared to the mouth and more accurate to classify neutral faces with fixation on the mouth compared to between the eyes. In a similar experiment with faces that subtended approximately 12° of visual angle vertically, Kliemann et al.'s (2012) participants were more accurate in classifying neutral and happy faces with fixation on the mouth compared to between the eyes, yet were not more accurate in classifying fear when fixating between the eyes as compared to the mouth.

Do first saccades preferentially target diagnostic features?

Studies using the brief-fixation paradigm have shown that proportionately more gaze changes occur upward from initial fixation on the mouth than downward from initial fixation on either eye or on the midpoint between the eyes, and, importantly, that this effect is modified by the viewed emotion. Typically, more fixation changes occur upward from the mouth than downward from the eyes for fearful, angry and neutral faces, a bias that is markedly reduced or eliminated for happy faces (Boll & Gamer, 2014; Gamer & Büchel, 2009; Gamer, et al., 2013; Gamer, et al., 2010; Kliemann, et al., 2010; Scheller, et al., 2012). This effect is not always found for fearful faces, however (Kliemann, et al., 2012). The authors of these studies summarize the upward saccades from the mouth as 'toward the eyes' and the downward saccades from the eyes as 'toward the mouth', and they even sometimes claim that their findings show that people reflexively saccade toward diagnostic emotional facial features (see especially Gamer & Büchel, 2009; Gamer, et al., 2013; Scheller, et al., 2012). Yet we remain more cautious in our interpretation of these findings, for the following reasons. First, the proportion saccade measure has been used to compare saccade direction at a very coarse level (up versus down) rather than to provide a measure of whether eye movements are toward specific facial features. Second, previous studies only compared the proportion of saccades downward from the eyes with the proportion of saccades upward from the mouth. A problem with this comparison is that the vertical distance from the center of the mouth to the center of the face is larger than the vertical distance from either eye, or from the midpoint between the eyes, to the center of the face, and there is a strong tendency for first saccades to be to the geometric center of scenes or configurations (e.g., Bindemann, Scheepers, Ferguson, & Burton, 2010; Findlay, 1982; P. Y. He & Kowler, 1989; Tatler, 2007), including faces (Bindemann, et al., 2009). Thus, the previously reported findings of proportionately more saccades upward from the mouth than downward from the eyes might be confounded by this “center-of-gravity effect”, though that effect would not fully explain the important additional finding that the difference in the proportion of upwards versus downward saccades varies as a function of the expressed emotion.

The Present Study

Using the brief-fixation technique across two experiments, we first tested whether the ability to identify facially expressed emotions varies according to the feature fixated. The fixation locations included the left and right eyes and the center of the mouth, but going beyond earlier research, also included the center of the brow and the left and right cheeks, as shown in Figure 1. (Note that ‘left’ and ‘right’ in this context refer to the left and right sides of the image.) The eyes themselves were chosen as fixation locations rather than the more commonly used midpoint between the eyes, for we were interested in investigating foveal vs. extrafoveal processing of the eyes specifically, given previous evidence that eyes are informative for the recognition of fear (Adolphs, et al., 2005; Smith, et al., 2005). The brow region was selected for being more informative about anger (Smith, et al., 2005) and, under our stimulus presentation conditions, was located approximately 1.3° above the midpoint between the eyes. The cheek locations were chosen as relatively uninformative parts of the face (compared to the eyes, mouth and brow) that were the same distance from the central point as the more informative features. For each face image, the positions of these fixation locations were equidistant from the center of the face (a point on the nose, but higher than the tip of the nose).

We predicted that emotion classification performance would be improved when the key emotion-distinguishing facial feature was aligned with fixation (and thus projected to the participant's fovea), relative to other facial features that are less or not informative of the expressed emotion, particularly when the emotion-relevant information is in the medium-to-high spatial frequencies. We also predicted that these effects would depend on what emotions and response options are pitted against each other, given the relative probabilities of confusions between different facially expressed emotions (e.g., Calder et al., 2000; Ekman & Friesen, 1976; Woodworth, 1938) and the finding that observers make use of different visual information from expressive faces depending on the combination of emotions and response options presented to them (Smith & Merlusca, 2014).

Figure 1. An example face image used in the experiments (from the Langner et al., 2010, database), overlaid, for illustrative purposes, with red (dark grey) crosses to mark the possible fixation locations, i.e., the locations on the face that were aligned with the fixation cross (which participants had to fixate in order for the face to appear). See the online article for the color version of this figure.

We also tested whether reflexive first saccades from initial fixation on the face (i.e., saccades triggered by the face stimulus but occurring after stimulus offset) preferentially target emotion-distinguishing facial features. To that end, we introduced a saccade projection measure, calculated by projecting the vector of a saccade on to the vectors from the initial fixation location to each of the other fixation locations (now regarded as saccade target locations), corrected for the length of the target vector (given that the target locations vary in distance from a given fixation location).

Experiment 1

In this experiment participants were presented with angry, fearful, happy, and emotionally neutral faces, replicating the combination of emotions used by most previous studies that have employed the brief-fixation paradigm. A novel contribution of the present experiment was the inclusion of a fixation location at the central brow, which contains medium-to-high spatial frequency information diagnostic of expressions of anger and sadness (Smith, et al., 2005), leading to the prediction that classification of angry faces would be improved when initial fixation was on the brow. We also predicted improved classification of fearful expressions when one or other of the eyes was aligned with fixation, given the diagnostic nature of the medium-to-high spatial frequency information in this region (Smith, et al., 2005; Smith & Merlusca, 2014). We further tested whether, as previously reported, first saccades after face offset would preferentially target the (now absent) upper face more than the lower face, and whether this effect would vary as a function of the emotion expressed on the face.

Method

Participants. Thirty students (22 female) aged 18-28 (*M* = 20.3 years, *SD* = 2.0) participated. All participants had normal or corrected-to-normal vision (assessed through self-report). All participants provided informed consent and received either course credit (Psychology students) or £6 (non-Psychology students) for their participation. The study was approved by the Ethics Committee of the Department of Psychology, Durham University.

We initially aimed for a sample size larger than those of most previous studies using the brief-fixation paradigm, which ranged from 15 (Neath & Itier, 2014) to 24 (Gamer & Büchel, 2009; Scheller, et al., 2012) for within-participants comparisons. We used the ESCI software (Cumming & Calin-Jageman, 2016; https://thenewstatistics.com/itns/esci/) to check whether our sample size of 30 provided acceptable precision for paired-samples t-tests, given that our main contrasts of interest were planned comparisons (e.g., greater accuracy for fixation on the brow relative to the cheeks of angry faces). The ESCI software computes the required *N* for a given target 95% confidence interval (CI) 'MoE' (halfwidth of the CI around the effect size). For a typical correlation between repeated measures, $\rho = 0.7$, the required *N* for a target MoE of 0.4 on average is 17 and with 99% assurance is 28. With $\rho = 0.6$, the required N for a target MoE of 0.4 on average is 22 and with 99% assurance is 35, whereas with $\rho = 0.8$, the required N for a target MoE of 0.4 on average is 13 and with 99% assurance is 21. Dropping the target MoE to 0.3, with $\rho = 0.7$, increased the required *N* to 29 on average and to 43 with 99% assurance.

Apparatus. Visual stimuli were presented on a CRT monitor with a viewable screen size of 340mm (width) × 245mm (height), a 1280 (width) × 960 (height) pixel resolution and 85 Hz refresh-rate. The participants were seated directly in front of the monitor with their head position controlled by a head and chin rest such that the viewing distance from the monitor screen was 48 cm. Graphics output was controlled by a Cambridge Research Systems (CRS) ViSaGe MKII Stimulus Generator. The display was gamma corrected (linearized) from measurements made with a CRS OptiCAL. The experiment was executed and controlled using the Matlab® programming language and CRS Toolbox functions. To control stimulus presentation and to measure gaze behavior, a CRS High-Speed Video Eye-Tracker was used, with a sampling-rate of 250 Hz. The eye-tracker was calibrated to each participant's right eye, but viewing was binocular.

Calibration was done using a nine-point automated calibration accuracy test. On each trial, participants made their responses by pressing one of the four outer buttons on a Cedrus RB-530 response pad.

Stimuli. Face images of 24 individuals (12 female, 12 male) were selected from the Radboud face database (http://www.socsci.ru.nl:8180/RaFD2/RaFD?p=main; Langner et al., 2010). All faces were of Caucasian adults with full frontal pose and gaze. Each individual was presented in each of 4 expressions: angry, fearful, happy and emotionally neutral. The 24 identities were selected from the larger set of 39 identities such that (a) there were equal numbers of females and males, and, based on the data from Langner et al.'s (2010) validation study, (b) more than 60% of participants labeled each of the 4 expressions posed by that person with the target emotion, and (c) these percentage agreement rates were balanced across emotions as much as possible. Although the percentage agreement did not differ substantially between the selected angry, fearful and neutral expressions (at 90.7%, 87.7% and 92.3%, respectively, all corrected *p*s > .1), the very high agreement rates for happy expressions (98.5% for the selected set, 98.9% for the larger set) prevented selection of a set of identities with equivalent agreement rates across all 4 emotions. All faces had been spatially aligned by Langner et al. (2010). All selected images were displayed in color. We cropped the images to 384 pixels (width) × 576 pixels (height), so that the face took up more of the image than in the original image set. The face images subtended 12.2° (width) × 17.3° (height) of visual angle at the viewing distance of 48 cm. On average, the faces within these images subtended 12.8° visual angle vertically, measured from the top of the forehead (at the hairline) to the tip of the chin, which is equivalent to the angular distance of another's face at the border between what Hall (1966) called close and far personal distances (Atkinson & Smithson, 2013).

Procedure. Participants were presented with 4 blocks of 96 trials (i.e., 384 trials in total).

The stimuli were randomly ordered across the 4 blocks, with a new random order for each participant. The blocks were separated by a self-paced break, and the eye-tracker was recalibrated at the beginning of each block. On each trial, the face image was preceded by a fixation cross, which could appear at any one of 25 locations at or near the center of the screen in order to make both the screen location of the fixation cross and the to-be-fixated facial feature unpredictable. These 25 possible locations for the fixation cross were at 0, 25, or 50 pixels left or right and up or down from the center of the screen. These fixation-cross positions were randomly ordered across trials. The faces were presented such that, on any given trial, one of 4 main locations on the face was aligned with the fixation cross: the center of each eye (left or right), the center of the mouth, the center of the brow, and the cheeks (left or right - directly below the eye). Henceforth we refer to these locations on the face as fixation locations (see Figure 1). Over the course of the 384 trials, each of the 96 images (4 expressions × 24 identities) was presented with each of the 4 facial locations aligned with the fixation cross. The face images were presented for approximately 82 ms (7 monitor refreshes), after which the screen remained blank (grey) until 500 ms after the participant made his or her manual response, when the fixation cross for the next trial appeared. A face image was presented only after the participant had been fixating within a radius of 15 pixels from the fixation cross for 3 consecutive eye-tracker samples. Participants were required to categorize faces as happy, angry, fearful or emotionally neutral, by pressing one of 4 buttons on a response box as quickly but as accurately as possible. The experiment was run in a darkened cubicle.

Analyses. Accuracy of emotion classification performance was assessed using unbiased hit rates (Wagner, 1993)[4], calculated for each emotion at each fixation location. The unbiased hit rate was chosen over standard hit rates because the frequency with which participants used the different response labels in each experiment was not equal (as reported in the Results sections for each experiment). To facilitate direct comparison with previously published studies, we report in the Supplementary Materials the same set of analyses using the standard proportion correct values. It is notable that these two sets of accuracy analyses produced similar results. Trials with response times less than or equal to 200 ms were treated as short RT outliers and were omitted from analyses of accuracy and response time. Median RTs were analyzed for correct responses only. Repeated-measures analyses of variance (ANOVAs) were conducted separately for each of these accuracy and RT measures, with the within-participant variables emotion (angry, fearful, happy, neutral) and fixation location (eyes, brow, mouth, cheeks).

The eye-tracking data were processed in Matlab using custom-made code. The eye-position data for each trial were first pre-processed to remove blinks, with the blink period defined as starting 5 samples (20ms) before the start of a segment of samples without eye-position data and ending 5 samples after the end of that segment of untracked samples.

A Savitzky-Golay smoothing filter was then applied to the trial-specific eye-position data, using the ‘sgolayfilt’ Matlab function, with second-order polynomials and a filter length of 9 samples (36 ms, i.e., just under twice the minimum saccade length of 5 samples/20 ms), as recommended by Nyström and Holmqvist (2010). Saccades were then identified as changes in eye position with a minimum velocity of 20^{o} per second and a minimum duration of 5 samples (20 ms), using either vertical or horizontal angular velocities, subsequent to the application of the same Savitzky-Golay smoothing filter to the velocity data. (In cases where both the vertical and horizontal angular velocities delineated the same saccade, the parameters for the saccade identified by the vertical angular velocity, such as the start and end times and positions, were selected for further processing and analysis.)

Following Gamer and colleagues (Boll & Gamer, 2014; Gamer & Büchel, 2009; Gamer, et al., 2013; Scheller, et al., 2012), we selected for analysis only the first saccades from enforced fixation on the face that occurred within 1000 ms of face offset and whose amplitude exceeded 1^{o} of visual angle (here referred to as ‘valid’ saccades). We subsequently calculated saccade projection measures as estimates of the paths of the reflexive saccades toward each of the face locations of interest (i.e., the fixation locations depicted in Figure 1, now regarded as target locations). First, six vectors were calculated from the starting coordinates of each valid saccade to the coordinates of the target locations. Then, the dot products of these possible saccade path vectors and the actual saccade vector were calculated and normalized to the magnitude of the saccade path vectors. This measure represents the similarity between the reflexive saccade path and the possible saccade path vectors. Identical vectors would produce a value of 1, whereas a saccade in exactly the same direction as a possible saccade path to a target but overshooting that target by 50% would produce a value of 1.5, for example. These normalized saccade projection measures were compared across experimental conditions using ANOVAs. We also supplement these analyses of the first-saccade trajectories with analyses comparing the proportions of first saccades upward from initial fixation on the lower-face locations (center of the mouth, left and right cheeks) with the proportions of first saccades downward from initial fixation on the upper-face locations (brow, left and right eyes). These analyses are reported in the Supplementary Materials, allowing the interested reader to compare directly the results of these analyses with those from previously published studies that used the same brief-fixation paradigm, all of which have used this alternative saccade measure. Frequencies of and mean latencies for reflexive saccades were also analyzed using ANOVAs, the results of which are also reported in the Supplementary Materials.

Statistical analyses were performed in JASP (version 0.10.2; https://jasp-stats.org/). Greenhouse-Geisser correction for degrees of freedom was used for all ANOVA main effects and interactions for which Mauchly’s test of sphericity had $p < .05$. All posthoc tests (pairwise and planned comparisons) were corrected for the relevant number of multiple comparisons using the Bonferroni-Holm method (corrected p-values are reported). Where the data failed to meet the normality assumption required for t-tests (as indicated by Shapiro-Wilk tests, $p < .05$), non-parametric Wilcoxon signed ranks tests were used. For ANOVAs, we report both the partial eta squared ($\eta_p{}^2$) and generalized eta squared ($\eta_G{}^2$) measures of effect size (Bakeman, 2005; Lakens, 2013; Olejnik & Algina, 2003); for t-tests and Wilcoxon tests, we report Hedges' *g* with 95% CI calculated using 'bias-corrected and accelerated' bootstrapping with the BootES package for R (Kirby & Gerlanc, 2013). (In the Supplementary Materials we also report the corresponding Cohen's d_z or matched-rank biserial correlations, with 95% CIs, as output by JASP.)

Results

For raw data and associated code, see Atkinson and Smithson (2019). The data for 10 trials (0.0868% of the total number of trials across all participants) were excluded from the accuracy and RT analyses; 6 because the RTs were less than or equal to 200 ms and 4 because the participant had pressed a key that was not one of the indicated response keys. A chi-squared test revealed considerable differences in the overall frequencies with which participants used the different response labels (irrespective of whether they were correctly applied), $\chi^2 = 99.5$, $p < .001$. Participants selected 'angry' rather less often (total = 2446 trials, participant $M = 81.5$, $SD = 16.2$) than they did the other 3 emotions ('fearful': total = 3156 trials, participant $M = 105.2$ trials, $SD = 10.1$; 'happy': total = 2879, participant $M = 96.0$, $SD = 2.8$; 'neutral': total = 3029, participant $M = 101.0$, $SD = 15.3$) and selected 'fearful' more often than they did 'happy' as well as 'angry'. This result motivated us to use the unbiased hit rates as a measure of emotion classification accuracy instead of standard hit rates.

We next performed an initial check for whether emotion classification performance varied as a function of the side of the face to which fixation on an eye or cheek was enforced. To this end, repeated-measures ANOVAs were conducted on the accuracy and RT measures, with the within-participant variables emotion (angry, fearful, happy, neutral), fixation location (eyes, cheeks), and side of face (left, right). There were no significant main effects of side of face and none of the interactions involving side of face was significant (all *p*s > .1). All subsequent analyses were therefore performed without side of face as a factor.

Accuracy. The unbiased hit rates are summarized in Figure 2a. The ANOVA revealed a large main effect of emotion, $F(1.92, 55.72) = 64.7$, $p < .001$, $\eta_p^2 = .69$, $\eta_G^2 = .394$, and a small main effect of fixation location, $F(3, 87) = 2.54$, $p = .062$, $\eta_p^2 = .08$, $\eta_G^2 = .011$. Given that our predictions entailed that the effect of fixation location on emotion classification accuracy would differ as a function of the emotion, it is notable that there was an interaction between emotion and fixation location, $F(5.66, 164.19) = 3.29$, $p = .005$, $\eta_p^2 = .102$, $\eta_G^2 = .017$. To follow-up the interaction, simple main effects analyses were conducted to examine the effect of fixation location for each emotion separately.

For angry faces, there was a main effect of fixation location, $F(3, 87) = 3.59$, $p = .017$, $\eta_p^2 = .11$, $\eta_G^2 = .028$. Three one-tailed planned comparisons (brow > cheeks, brow > eyes, brow > mouth) revealed greater accuracy for expressions of anger when participants fixated the brow ($M = 0.712$, $SD = 0.155$) as compared to the cheeks ($M = 0.645$, $SD = 0.181$), $W = 359$, $p = .004$, $g = 0.42$, 95% CI [0.03, ∞], and eyes ($M = 0.66$, $SD = 0.174$), $W = 357$, $p = .001$, $g = 0.42$, 95% CI [0.0, ∞], but not mouth ($M = 0.705$, $SD = 0.183$), $t(29) = 0.33$, $p = .74$, $g = 0.06$, 95% CI [-0.25, ∞] (minimum Bonferroni-Holm adjusted $\alpha = .0167$). An examination of the confusion matrices in Figure 3 reveals that the improved ability to identify angry expressions when the brow was at fixation was principally associated with a reduction in the number of misclassifications of angry faces as neutral. Interestingly, there was also a small reduction in the number of misclassifications of angry faces as fearful when initial fixation was on a lower-face feature (one or other cheek and especially the mouth) than when initial fixation was on an upper-face feature (the brow or an eye).

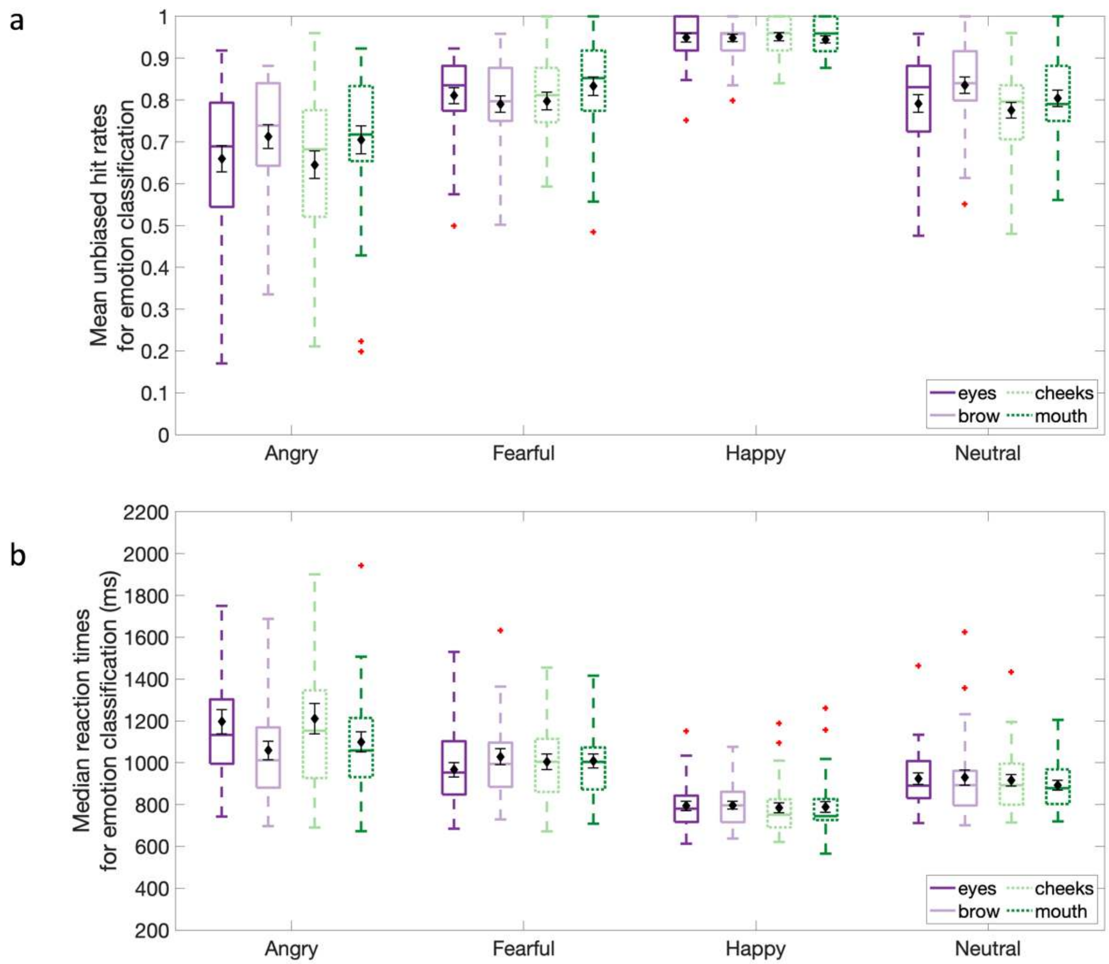

Figure 2. Emotion classification accuracy (a: mean unbiased hit rates) and median response times (b) as a function of emotion category and fixation location in Experiment 1. On each box, the central horizontal line indicates the median value across participants and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers. (Outliers, indicated with a '+', were defined as individual participant mean values more than 2 times the interquartile range away from the top or bottom of the box, and were not excluded from the data analyses.) Overlaid on the boxes are the mean values across participants, indicated by black diamonds, with error bars indicating the SEMs. See the online article for the color version of this figure.

For fearful faces, there was little effect of fixation location on unbiased hit rates, either as revealed by the main effect, $F(3, 87) = 2.1$, $p = .106$, $\eta_p^2 = .067$, $\eta_G^2 = .021$, or by three one-tailed planned comparisons testing the prediction that participants would be more accurate in classifying fearful faces when fixating the eyes (eyes > cheeks, eyes > mouth, eyes > brow; all *p*s > .05). An examination of the full pairwise comparisons revealed a small trend for greater unbiased hit rates when participants fixated the mouth ($M = 0.833$, $SD = 0.121$) than when they fixated the brow of fearful faces ($M = 0.79$, $SD = 0.107$), $t(29) = 2.7$, $p = .011$, $g = 0.48$, 95% CI [0.11, 0.82] (two-tailed, minimum Bonferroni-Holm adjusted $\alpha = .0083$). An examination of the confusion matrices in Figure 3 reveals that the small improvement in the classification of fearful expressions for fixations on the mouth compared to fixations on the brow was principally associated with a reduction in the number of misclassifications of fearful faces as angry. More generally, fearful faces were less often misclassified as being angry when fixation was on the eyes or mouth than when fixation was on the brow or cheeks.

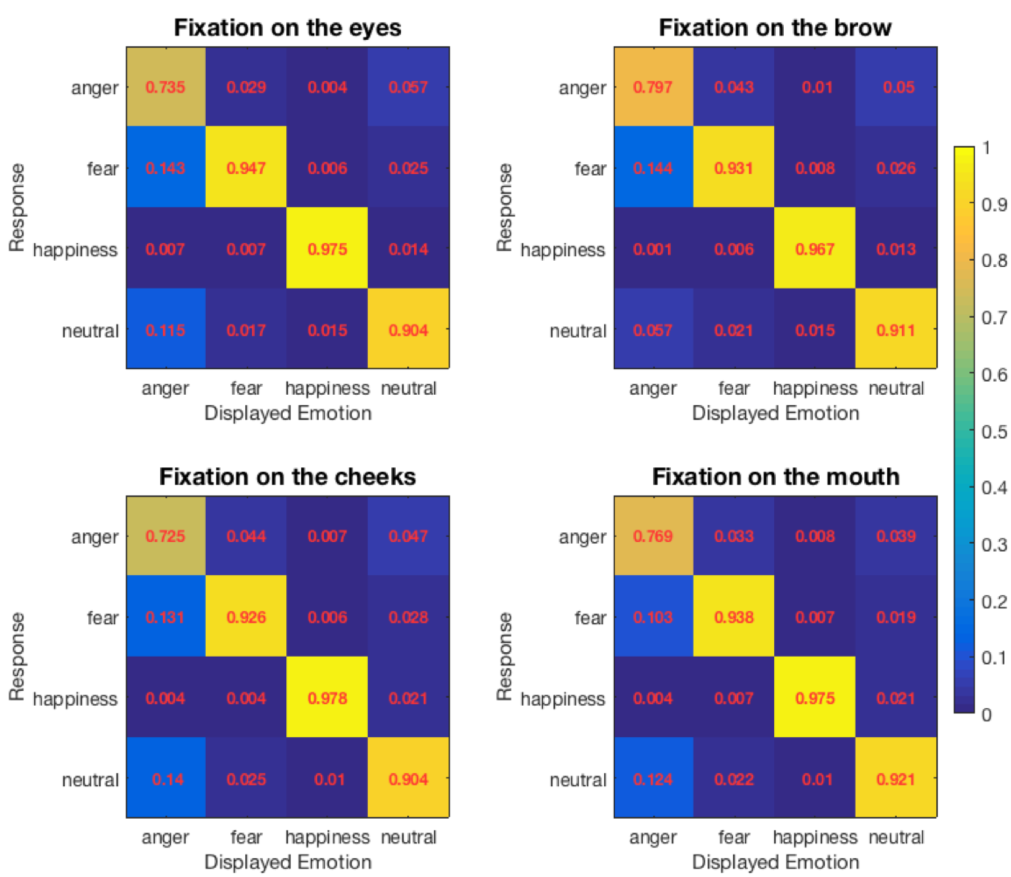

Figure 3. Confusion matrices for Experiment 1. The color scale and the corresponding numbers in the cells indicate the mean proportion of times an emotion label was selected (y axis) for a given facially expressed emotion (x axis). See the online article for the color version of this figure.

For neutral faces, there was a medium-sized effect of fixation location on unbiased hit rates, $F(3, 87) = 3.63$, $p = .016$, $\eta_p^2 = .111$, $\eta_G^2 = .04$. Given that much of the central face is informative for classifying emotionally neutral expressions (Smith, et al., 2005; Smith & Merlusca, 2014), we did not have any specific predictions as to the effect of initial fixation location on neutral faces. We therefore conducted a full set of pairwise comparisons between fixation locations, which revealed that participants were more accurate in classifying neutral faces when fixating the brow ($M = 0.835$, $SD = 0.11$) as compared to the cheeks ($M = 0.775$, $SD = 0.104$), $t(29) = 3.03$, $p = .005$, $g = 0.54$, 95% CI [0.12, 0.93] (two-tailed, minimum Bonferroni-Holm adjusted $\alpha = .0083$), which was associated principally with fewer misclassifications of neutral faces as happy (Figure 3). For happy faces, there was no main effect of fixation location on unbiased hit rates ($F < 1$, $p > .9$).

Response times. The RT data are summarized in Figure 2b. There was a large main effect of emotion, $F(1.88, 54.46) = 43.12$, $p < .001$, $\eta_p^2 = .598$, $\eta_G^2 = .278$, and all pairwise comparisons of the 4 emotions were significant (all ps < .004, minimum Bonferroni-Holm adjusted $\alpha = .0042$), reflecting shorter RTs for happy faces compared to the other three emotions, for neutral faces compared to angry and fearful faces, and for fearful compared to angry faces. There was also a small influence of fixation location on RTs, $F(3, 87) = 2.38$, $p = .075$, $\eta^2 = .076$, $\eta_G^2 = .004$. Importantly, these main effects were modified by an interaction, $F(3.93, 113.83) = 4.22$, $p = .003$, $\eta_p^2 = .127$, $\eta_G^2 = .024$.

Simple main effects analyses revealed a large effect of fixation location for angry faces, $F(2.15, 62.42) = 5.9$, $p = .004$, $\eta_p^2 = .169$, $\eta_G^2 = .042$. Participants were faster to classify angry faces correctly when fixating the brow (*M* = 1059ms, *SD* = 240) as compared to the eyes (*M* = 1196ms, *SD* = 317), *W* = 368, *p* = .003, *g* = 0.54, 95% CI [0.26, ∞], and cheeks (*M* = 1211ms, *SD* = 397), *W* = 369, *p* = .002, *g* = 0.49, 95% CI [0.15, ∞], as predicted, but not as compared to the mouth (*M* = 1100ms, *SD* = 264), *t*(29) = 1.29, *p* = .1, *g* = 0.23, 95% CI [-0.1, ∞] (one-tailed planned comparisons; minimum Bonferroni-Holm adjusted α = .0167).

For fearful faces, fixation location had a small effect on RTs, *F*(3, 87) = 2.37, *p* = .076, η_p^2 = .075, η_G^2 = .014. RTs were shorter when initial fixation was on an eye of fearful faces (*M* = 966ms, *SD* = 188) than when initial fixation was on the brow (*M* = 1029ms, *SD* = 207), *t*(29) = 2.96, *p* = .003, *g* = 0.53, 95% CI [0.17, ∞], but not compared to when initial fixation was on a cheek (*M* = 1005ms, *SD* = 204), *W* = 295, *p* = .1, *g* = 0.28, 95% CI [-0.01, ∞], or the mouth (*M* = 1009ms, *SD* = 182), *t*(29) = 1.62, *p* = .058, *g* = 0.29, 95% CI [-0.02, ∞] (one-tailed planned comparisons; minimum Bonferroni-Holm adjusted α = .0167). For happy and neutral faces there were no main effects of fixation location (both *F*s < 1, *p*s > .5).

Thus, the effects of initial fixation location on accuracy, reported in the previous section, do not reflect a speed-accuracy trade-off. That is, in none of the conditions in which participants were more accurate were they also slower; indeed, in the case of initial fixation to the brow of angry faces, participants tended to be both more accurate and faster to classify the emotion.

Saccade analyses. Five participants produced valid saccades (i.e., first saccades that occurred within 1000 ms of face offset and whose amplitude exceeded 1° of visual angle) on fewer than 20% of trials overall and so the data for these participants were excluded from all saccade analyses. For the remaining 25 participants, 54.7% of trials on average had a valid saccade (*SD* = 17.37, range: 25.5-86.5%). Normalized saccade projections with absolute values > 1.5 (excluding trajectories from the initial fixation location to itself) were classed as outliers, which effectively excluded saccades that ended close to or beyond the edge of the face. Such outliers accounted for 0.88% of the recorded measures and were excluded from the analyses.

Analyses of the proportions of saccades upward from fixation on the mouth vs. downward from fixation on the eyes or brow (see Supplementary Materials) replicated previous findings of more upward than downward saccades (Boll & Gamer, 2014; Gamer & Büchel, 2009; Gamer, et al., 2010; Kliemann, et al., 2012; Kliemann, et al., 2010; Scheller, et al., 2012), which here was evident for angry and neutral faces, less so for fearful faces, and not at all for happy faces (see Discussion for more details). There were not proportionately more reflexive saccades upwards from the cheeks than downwards from either the eyes or brow, however, for any of the 4 emotions.

Given that the proportion of saccades upwards vs. downwards is an imprecise measure of saccade direction, our main saccade analyses of interest used the normalized saccade projection measure. Using this measure, we tested the hypothesis that reflexive first saccades seek out (i.e., are more strongly in the direction of) emotion-informative features. To this end, we first performed a repeated-measures ANOVA on the normalized saccade projection measures with the data collapsed across initial fixation location. The directional strength of reflexive first saccades towards the 6 target locations, collapsed across fixation location, are summarized as a function of expressed emotion in Figure 4a. The normalized saccade projections varied as a function of target location, $F(1.89, 45.28) = 3.72$, $p = .034$, $\eta_p^2 = .134$, $\eta_G^2 = .069$, but not as a function of emotion or of their interaction (both Fs < 1.3, ps > .25). Regardless of starting location and emotion, first saccades from face offset were more strongly in the direction of the brow than the right eye, $t(24) = 4.99$, $p < .001$, $g = 0.97$, 95% CI [0.45, 1.43], and more strongly in the direction of the mouth than the right cheek, $t(24) = 4.73$, $p < .001$, $g = 0.92$, 95% CI [0.33, 1.55] and right eye, $t(24) = 2.9$, $p = .008$, $g = 0.56$, 95% CI [0.05, 1.05] (minimum Bonferroni-Holm adjusted $\alpha = .0033$).

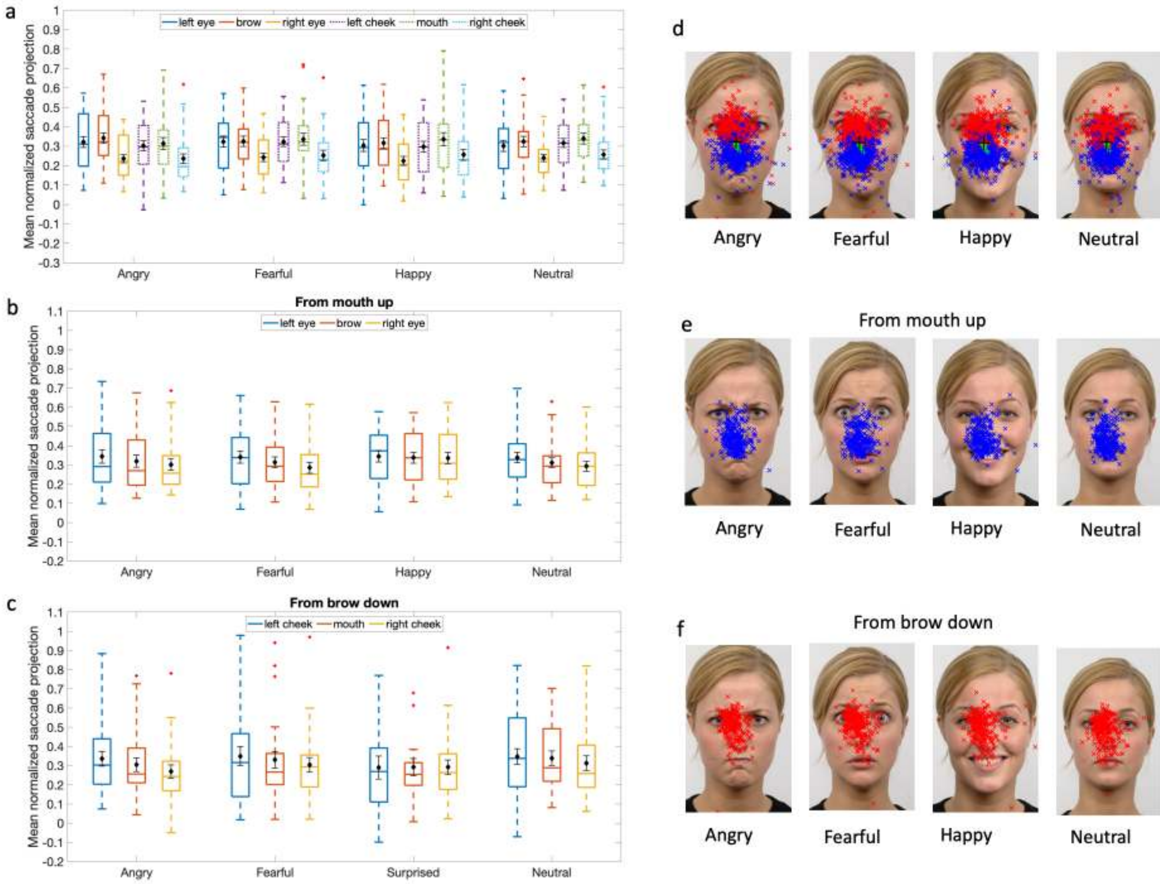

Figure 4. Mean normalized saccade projections as a function of the emotion expressed on the face and the target locations of interest (a – c) and saccade endpoints (d – f) for Experiment 1 (N = 25). The normalized saccade projection is a measure of the directional strength of the reflexive first saccades (executed after face offset) towards target locations of interest, in this case (a) to 6 target locations, collapsed across initial fixation location, (b) from the mouth center upwards towards the 3 upper-face target locations, and (c) from the central brow downwards towards the other 3 lower-face target locations. On each box, the central horizontal line indicates the median value across participants and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers. (Outliers, indicated with a '+', were defined as individual participant mean values more than 2 times the interquartile range away from the top or bottom of the box, and were not excluded from the data analyses.) Overlaid on the boxes are the mean values across participants, indicated by black diamonds, with error bars indicating the SEMs. (d) The endpoints for first saccades for each emotion, collapsed across participants and initial fixation location, plotted on example faces (which were not visible at the time of saccade execution). Blue crosses indicate endpoints of saccades from initial fixation on one of the lower-face features (mouth, cheeks) and red crosses indicate endpoints of saccades from initial fixation on one of the upper-face features (eyes, brow). The small green '+' indicate individual subject mean coordinates for their saccade endpoints, whereas the larger black '+' indicate the mean coordinates for first saccade endpoints, averaged over all trials for all subjects. The endpoints for first saccades upwards from fixation on the center of the mouth and downwards from the central brow are plotted separately in (e) and (f).

The example face images in (d)-(f) are from the Langner et al. (2010) database. See the online article for the color version of this figure.

Next, given that we observed proportionately more reflexive first saccades upwards from the mouth than downwards from the eyes or brow, we tested whether those upward-directed saccades from the mouth more strongly targeted one or other (or both) of the eyes or the brow and whether this varied as a function of the emotional expression. The relevant normalized saccade projection data are summarized in Figure 4b. There was a main effect of target location, $F(1.01, 23.13) = 4.7$, $p = .041$, $\eta_p^2 = .17$, $\eta_G^2 = .012$, and a marginally significant interaction, $F(2.43, 55.97) = 2.47$, $p = .083$, $\eta_p^2 = .097$, $\eta_G^2 = .002$, but no main effect of emotion ($F < 1$, $p >$ .45.). Collapsed across emotion, reflexive first saccades upwards from the mouth were more strongly in the direction of the left eye ($M = 0.342$, $SD = 0.141$) than of the brow ($M = 0.32$, $SD =$ 0.126), $t(24) = 2.5$, $p = .02$, $g = 0.48$, 95% CI [0.06, 0.95] and right eye ($M = 0.303$, $SD = 0.127$), $t(24) = 2.24$, $p = .034$, $g = 0.43$, 95% CI [0.02, 0.88], and marginally more strongly in the direction of the brow than of the right eye, $t(24) = 1.97$, $p = .061$, $g = 0.38$, 95% CI [-0.05, 0.81]. This reflected a leftward bias for the upward-directed saccades leaving the mouth, also indicated by a linear decrease in the normalized saccade projection values from left to right across the 3 target locations, $t(24) = -3.06$, $p = .004$ (polynomial contrast). Simple main effects analyses revealed that this linear trend in normalized saccade projection values for upward-directed saccades leaving the mouth was evident for the angry, fearful and neutral faces but not for the happy faces, for which there were no differences across target locations (main effect of target location: $F < 0.3$, $p > .6$; polynomial contrasts: $|t|$s < 0.75, ps $> .45$).

Finally, we tested whether the downward-directed saccades from the brow more strongly targeted the mouth or one or other (or both) of the cheeks and whether this varied as a function of the emotional expression.[5] The relevant normalized saccade projection data are summarized in

Figure 4c. There was a main effect of target location, $F(1.01, 23.2) = 9.53$, $p = .005$, $\eta_p^2 = .293$, $\eta_G^2 = .034$, but no main effect of emotion or interaction (*F*s < 1.9, *p*s > .15). Collapsed across emotion, reflexive first saccades downwards from the brow were more strongly in the direction of the left cheek ($M = 0.329$, $SD = 0.13$) than of the mouth ($M = 0.299$, $SD = 0.126$), $t(24) = 2.79$, $p = .01$, $g = 0.54$, 95% CI [0.11, 1.02] and right cheek ($M = 0.265$, $SD = 0.137$), $t(24) = 3.11$, $p = .005$, $g = 0.6$, 95% CI [0.14, 1.07], and more strongly in the direction of the mouth than of the right cheek, $t(24) = 3.4$, $p = .002$, $g = 0.66$, 95% CI [0.21, 1.16]. This reflected a leftward bias for the downward-directed saccades leaving the brow, also indicated by a linear decrease in the normalized saccade projection values from left to right across the 3 target locations, $t(24) = -4.36$, $p < .001$ (polynomial contrast).

Discussion

As predicted, allowing participants a single fixation on only the central brow region of the face, with the rest of the face thus projecting to extrafoveal retina, improved emotion classification performance for angry faces, relative to single fixations on a cheek or eye (but not mouth) of those angry faces. This improvement was evident in both greater accuracy and shorter RTs. These novel findings are consistent with previous work showing the importance of high spatial-frequency information from the brow in allowing observers to distinguish angry expressions from expressions of other basic emotions (Smith, et al., 2005).

We did not find a similar improvement in the classification of fearful faces when fixation was forced on either the left or right eye, however, as has been reported in one previous study for accuracy (Kliemann, et al., 2010) and two for RTs (Gamer, et al., 2010; Kliemann, et al., 2010). Other studies have similarly found no statistically significant effect on fear classification performance for fixation on the eyes, as indicated by accuracy (Gamer & Büchel, 2009; Gamer, et al., 2010; Neath & Itier, 2014; Scheller, et al., 2012) or both accuracy and RTs (Kliemann, et al., 2012). Our and these other null results indicate that foveal processing of the eyes in fearful faces (and thus the extraction of high spatial-frequency information) provides little or no discernable benefit, relative to extrafoveal processing of those eyes (and thus the extraction of relatively low spatial-frequencies), in allowing observers to discriminate fearful expressions from at least angry, happy and neutral expressions. Interestingly, though, we found that fixation on the mouth improved classification accuracy for fearful faces relative to fixation on the brow, and that this was principally associated with a reduction in the number of misclassifications of fearful faces as angry. This result is consistent with Smith and Merlusca's (2014) finding that, in free-viewing tasks, part of the information that observers rely on for the successful classification of fearful compared to angry and emotionally neutral faces is mid-range spatial frequency information in the central mouth region, whereas when asked to distinguish only between fearful and neutral faces, mid-to-high spatial frequencies in the central brow region constitute another source of information upon which they rely.

We replicated previous findings (Boll & Gamer, 2014; Gamer & Büchel, 2009; Gamer, et al., 2013; Gamer, et al., 2010; Kliemann, et al., 2010; Scheller, et al., 2012) of more saccades upward from the mouth than downward from the eyes (or from between the eyes), using the same emotions as in those studies (anger, fear, happiness and neutral; or, for some studies, just fear, happiness and neutral). We found that this effect was evident for angry and neutral faces, less so for fearful faces, and not at all for happy faces. Nonetheless, there were not proportionately more reflexive saccades upwards from the cheeks (which can be considered as lower-face locations) than downwards from either the eyes or brow, for any of the four emotions. Interestingly, for fearful faces, proportionately fewer reflexive saccades were made downwards from the eyes than from the brow, perhaps reflecting differences in the informativeness of these features for fear.

Although we found that proportionately more reflexive first saccades were made upwards from the mouth than downwards from the brow (or eyes), our saccade projection analyses did not indicate that those saccades preferentially target emotion-distinguishing facial features. Instead, regardless of the emotion shown on the face, the reflexive first saccades exhibited a directional bias away from features on the right to features in the center and even more so to features on the left side of the face. Averaged across starting location and emotion, first saccades from face offset were more strongly in the direction of the central face features (brow and mouth) than features on the right of the face (right eye and cheek). The reflexive saccades that were directed upward from the center of the mouth exhibited a leftward bias, that is, were more strongly in the direction of the central brow than the right eye and more strongly in the direction of the right eye than the brow. This was the case for angry, fearful and neutral faces, but not for happy faces. The reflexive saccades directed downwards from fixation on the brow showed a similar leftward bias for the lower-face features (left cheek, mouth, right cheek), and was evident for all four emotions.

Experiment 2

In this experiment participants were presented with angry, fearful, surprised, and emotionally neutral faces. Based on the results of Experiment 1, we predicted improved classification of angry faces when initial fixation was on the brow. We also predicted that initial fixation on the mouth would elicit improved classification performance for fearful and surprised faces, given that it is the mouth region that principally distinguishes fearful from surprised as well as from neutral and angry expressions, whereas the eye region does not distinguish between prototypical fearful and surprised expressions (Du, Tao, & Martinez, 2014; Roy-Charland, Perron, Beaudry, & Eady, 2014; Smith, et al., 2005). As with Experiment 1, we also tested whether reflexive first saccades would preferentially target emotion-distinguishing facial features.

Method

Participants. Thirty students (22 female) aged 18-23 (*M* = 19.7 years, *SD* = 1.3) participated; none had taken part in Experiment 1. All participants had normal or corrected-to-normal vision (assessed through self-report). All participants provided informed consent and received either course credit (Psychology students) or £6 (non-Psychology students) for their participation. The study was approved by the Ethics Committee of the Department of Psychology, Durham University.

Stimuli. The stimuli were identical to those used in Experiment 1, except that surprised faces were substituted for happy faces; that is, participants were presented with faces expressing surprise, anger, fear, or no emotion (neutral). The surprised faces were selected from the same database as angry, fearful and neutral faces.

Apparatus. The apparatus was identical to that for Experiment 1.

Procedure. The procedure was identical to that for Experiment 1, except that participants were required to categorize faces as surprised (instead of happy), angry, fearful or neutral.

Analyses. The analyses performed and criteria adopted were the same as those in Experiment 1.

Results

For raw data and associated code, see Atkinson and Smithson (2019). The data for 12 trials (0.104% of the total number of trials across all participants) were excluded from the accuracy and RT analyses; 9 because the RTs were less than or equal to 200ms and 3 because the participant had pressed a key that was not one of the indicated response keys. A chi-squared test revealed differences in the overall frequencies with which participants used the different response labels (irrespective of whether they were correctly applied), $\chi^2 = 12.8$, $p < .01$. Participants selected ‘surprised’ (total = 2994 trials, participant *M* = 99.8, *SD* = 11.3) and ‘neutral’ (total =

2936 trials, participant *M* = 97.9, *SD* = 10.4) more often than they selected 'fearful' (total = 2751 trials, participant *M* = 91.7, *SD* = 12.1) and 'angry' (total = 2827 trials, participant *M* = 94.2, *SD* = 11.8). This result motivated us to use the unbiased hit rates as a measure of emotion classification accuracy instead of standard hit rates.

We next performed an initial check for whether emotion classification performance varied as a function of the side of the face to which fixation on an eye or cheek was enforced. To this end, repeated-measures ANOVAs were conducted on the accuracy and RT measures, with the within-participant variables emotion (angry, fearful, surprised, neutral), fixation location (eyes, cheeks), and side of face (left, right). There were no significant main effects of side of face and none of the interactions involving side of face was significant (all *p*s > .1). All subsequent analyses were therefore performed without side of face as a factor.

Accuracy. The unbiased hit rates are summarized in Figure 5a. The ANOVA revealed a main effect of emotion, $F(1.45, 41.95) = 41.86$, $p < .001$, $\eta_p^2 = .591$, $\eta_G^2 = .297$, and a main effect of fixation location, $F(3, 87) = 3.55$, $p = .018$, $\eta_p^2 = .109$, $\eta_G^2 = .019$, the latter reflecting larger unbiased hit rates for initial fixation on the mouth than on the cheeks ($p = .004$; minimum Bonferroni-Holm adjusted $\alpha = .0083$). Crucially, given our predictions entailing that the effect of fixation location on emotion classification accuracy would differ as a function of the emotion, these main effects were modified by an interaction, $F(4.57, 132.56) = 6.63$, $p < .001$, $\eta_p^2 = .186$, $\eta_G^2 = .039$. To follow-up the interaction, simple main effects analyses were conducted to examine the effect of fixation location for each emotion separately.

For angry faces, there was evidence that fixation location had an effect on unbiased hit rates, $F(3, 87) = 2.71$, $p = .05$, $\eta_p^2 = .086$, $\eta_G^2 = .039$. Three one-tailed planned comparisons (brow > eyes, brow > cheeks, brow > mouth) revealed larger unbiased hit rates for expressions of anger when participants fixated the brow (*M* = 0.777, *SD* = 0.103) as compared to the cheeks (*M*

= 0.708, *SD* = 0.106), $t(29) = 2.59$, $p = .008$, $g = 0.46$, 95% CI [0.18, ∞], and eyes (*M* = 0.731, *SD* = 0.113), $t(29) = 2.1$, $p = .023$, $g = 0.37$, 95% CI [0.05, ∞], but not mouth (*M* = 0.745, *SD* = 0.124), $t(29) = 1.33$, $p = .097$, $g = 0.24$, 95% CI [-0.05, ∞] (minimum Bonferroni-Holm adjusted $\alpha = .0167$). An examination of the confusion matrices in Figure 6 reveals that the improved ability to identify angry expressions when the brow was at fixation was principally associated with a reduction in the number of misclassifications of angry faces as neutral, especially compared to when fixation was on the lower-face features (cheeks and mouth).

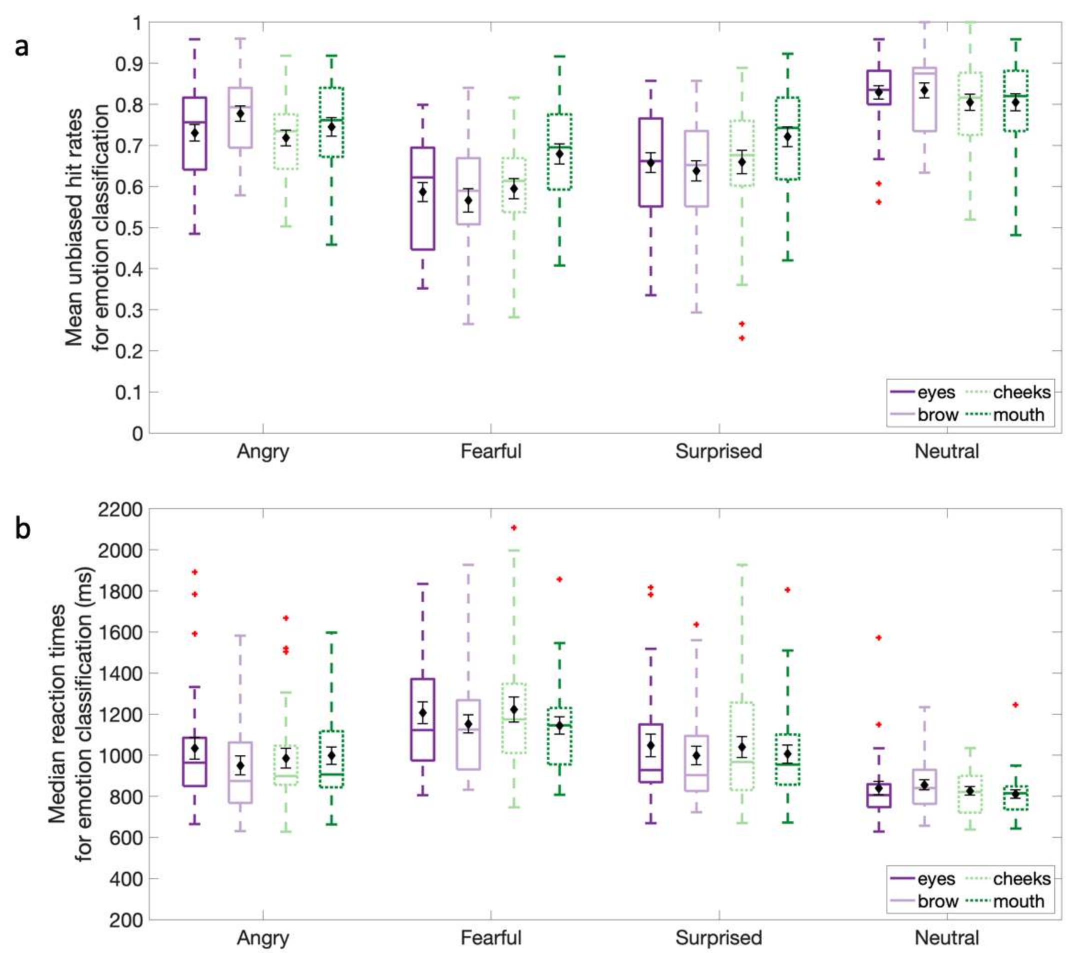

Figure 5. Emotion classification accuracy (a: mean unbiased hit rates) and median response times (b) as a function of emotion category and fixation location in Experiment 2. On each box, the central horizontal line indicates the median value across participants and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers. (Outliers, indicated with a '+', were defined as

individual participant mean values more than 2 times the interquartile range away from the top or bottom of the box, and were not excluded from the data analyses.) Overlaid on the boxes are the mean values across participants, indicated by black diamonds, with error bars indicating the SEMs. See the online article for the color version of this figure.

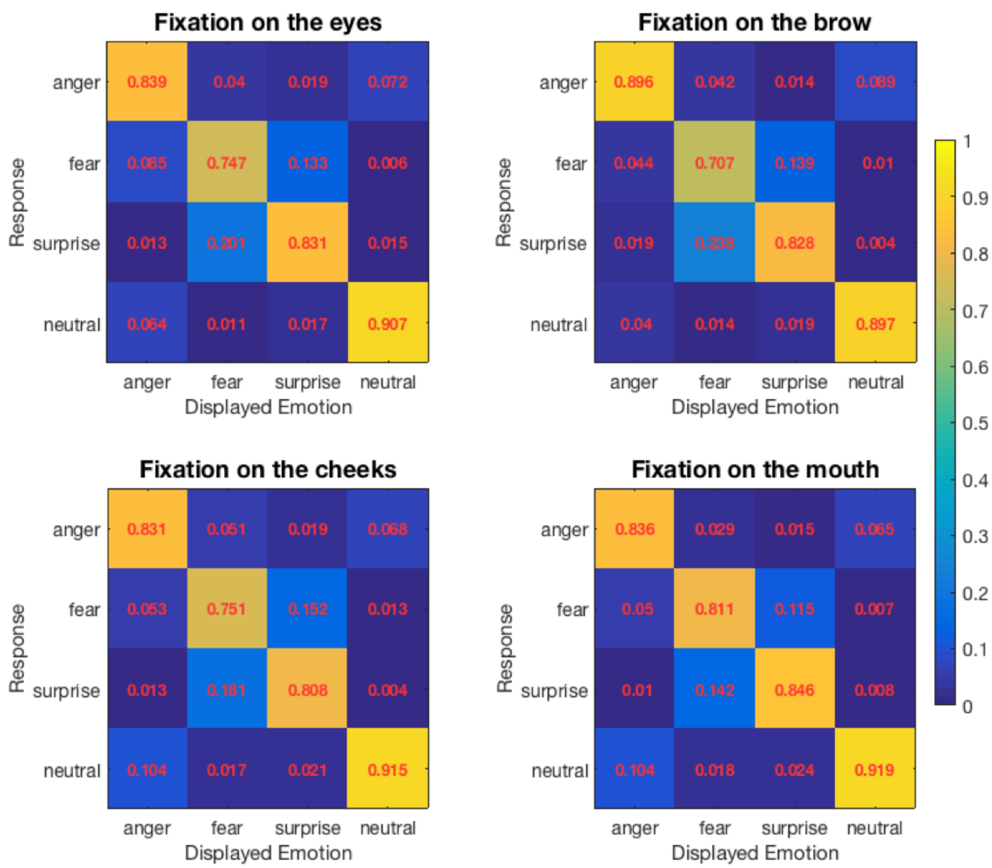

Figure 6. Confusion matrices for Experiment 2. The color scale and the corresponding numbers in the cells indicate the mean proportion of times an emotion label was selected (y axis) for a given facially expressed emotion (x axis). See the online article for the color version of this figure.

For fearful faces, there was a large main effect of fixation location, $F(3, 87) = 8.29$, $p < .001$, $\eta_p^2 = .222$, $\eta_G^2 = .092$. To follow up this main effect, we conducted three one-tailed planned comparisons to test the prediction that forcing fixation on the mouth would elicit greater classification accuracy for fearful faces, compared to fixation on the eyes, brow or cheeks (i.e., mouth > eyes, mouth > brow, mouth > cheeks). Recall that it is the mouth region that distinguishes fearful from neutral, angry and, particularly, surprised expressions, whereas the eye region does not distinguish between fearful and surprised expressions (Du, et al., 2014; Roy-Charland, et al., 2014; Smith, et al., 2005). These planned comparisons revealed larger unbiased hit rates for expressions of fear when participants fixated the mouth (*M* = 0.68, *SD* = 0.136) as compared to the eyes (*M* = 0.587, *SD* = 0.127), $t(29) = 4.07$, $p < .001$, $g = 0.72$, 95% CI [0.36, ∞], cheeks (*M* = 0.594, *SD* = 0.134), $t(29) = 3.85$, $p < .001$, $g = 0.68$, 95% CI [0.34, ∞], and brow (*M* = 0.566, *SD* = 0.155), $t(29) = 3.79$, $p < .001$, $g = 0.67$, 95% CI [0.34, ∞] (minimum Bonferroni-Holm adjusted $\alpha = .0167$). An examination of the confusion matrices in Figure 6 reveals that the improved ability to identify fearful expressions when the mouth was at fixation was principally associated with a reduction in the number of misclassifications of fearful faces as surprised, particularly compared to when fixation was on the upper-face features (eyes and brow) but also compared to when fixation was on the cheeks. There was also a reduction in the number of misclassifications of fearful faces as angry when fixation was on the mouth as compared to the other three fixation locations.

The same pattern held for the unbiased hit rates for surprised faces: a large main effect of fixation location, $F(3, 87) = 6.97$, $p < .001$, $\eta_p^2 = .194$, $\eta_G^2 = .05$, reflected greater accuracy for expressions of surprise when participants fixated the mouth (*M* = 0.721, *SD* = 0.133) as compared to the brow (*M* = 0.638, *SD* = 0.134), $t(29) = 3.84$, $p < .001$, $g = 0.68$, 95% CI [0.34, ∞], cheeks (*M* = 0.66, *SD* = 0.156), $t(29) = 3.1$, $p = .002$, $g = 0.55$, 95% CI [0.22, ∞], and eyes (*M* = 0.658, *SD* = 0.134), $t(29) = 3.17$, $p = .002$, $g = 0.56$, 95% CI [0.21, ∞] (one-tailed planned comparisons; minimum Bonferroni-Holm adjusted $\alpha = .0167$). The improved ability to classify surprised faces when the mouth was at fixation was principally associated with a reduction in the number of misclassifications of surprised faces as fearful (Figure 6). There was little effect of fixation location on unbiased hit rates for neutral faces, $F(3, 87) = 1.36$, $p = .26$, $\eta_p^2 = .045$, $\eta_G^2 = .018$.

Response times. The RT data are summarized in Figure 5b. There were main effects of emotion, $F(3, 87) = 34.38$, $p < .001$, $\eta_p^2 = .542$, $\eta_G^2 = .212$, and fixation location, $F(3, 87) = 3.72$, $p = .014$, $\eta_p^2 = .114$, $\eta_G^2 = .006$, but the Emotion × Fixation Location interaction effect was small and non-significant, $F(5.73, 166.25) = 1.33$, $p > .2$, $\eta_p^2 = .044$, $\eta_G^2 = .006$. Despite the nonsignificant interaction, our a priori hypotheses motivated planned comparisons. For angry faces, RTs were shorter for fixations on the brow (M = 950ms, SD = 251) than for fixations on the eyes (M = 1034ms, SD = 291), $t(29) = 3.12$, $p = .002$, $g = 0.56$, 95% CI [0.28, ∞], but not cheeks (M = 986ms, SD = 263), $t(29) = 1.4$, $p = .087$, $g = 0.25$, 95% CI [-0.04, ∞] or mouth (M = 998ms, SD = 229), $t(29) = 1.67$, $p = .053$, $g = 0.3$, 95% CI [-0.01, ∞]. For fearful faces, RTs were shorter for fixations on the mouth (M = 1145ms, SD = 229) than for fixations on the cheeks (M = 1223ms, SD = 332), $W = 328$, $p = .025$, $g = 0.44$, 95% CI [0.18, ∞], and marginally so compared to fixation on the eyes (M = 1206ms, SD = 292), $t(29) = 2.01$, $p = .027$, $g = 0.36$, 95% CI [0.06, ∞], but not brow (M = 1153ms, SD = 245), $t(29) = 0.27$, $p = .4$, $g = 0.05$, 95% CI [-0.25, ∞]. (All tests one-tailed; minimum Bonferroni-Holm adjusted $\alpha = .0167$.) There were no clear differences in RTs for surprised faces as a function of fixation location (all uncorrected ps > .1).

Thus, similar to Experiment 1, the effects of initial fixation location on accuracy do not reflect a speed-accuracy trade-off. Indeed, for angry and fearful faces, participants were not only more accurate but also faster to classify the emotion when initially fixating the distinguishing feature (brow or mouth, respectively) compared to other locations on the face.

Saccade analyses. Three participants produced valid first saccades (i.e., saccades that occurred within 1000 ms of face offset and whose amplitude exceeded 1° of visual angle) on fewer than 20% of trials overall and so the data for these participants were excluded from all saccade analyses. For the remaining 27 participants, 56.8% of trials on average had a valid saccade (SD = 17.34, range: 29.7-88.3%). Normalized saccade projections with absolute values >

1.5 (excluding trajectories from the initial fixation location to itself) were classed as outliers, which accounted for 1.36% of the recorded measures, and were excluded from the analyses.

Analyses of the proportions of reflexive saccades (see Supplementary Materials) replicated previous findings and our finding in Experiment 1 of more saccades upward from the mouth than downward from the eyes or brow. In the present experiment, however (in which only one of the four emotions was different, i.e., surprised instead of happy faces), we did not replicate the finding that this effect varies as a function of the expressed emotion.

As with Experiment 1, our main saccade analyses of interest used the normalized saccade projection measure to test the hypothesis that reflexive first saccades are more strongly in the direction of emotion-informative features. To this end, we first performed a repeated-measures ANOVA on the normalized saccade projection measures, with the data collapsed across initial fixation location (the data for which are summarized in Figure 7a). There was an influence of target location on the directional strength of first saccades, $F(2, 52.01) = 5.45$, $p = .007$, $\eta_p^2 = .173$, $\eta_G^2 = .076$, but no effect of emotion ($F < 1$, $p > .6$) and only little effect of their interaction, $F(6.1, 158.56) = 1.63$, $p = .142$, $\eta_p^2 = .059$, $\eta_G^2 = .005$. Regardless of starting location and emotion, first saccades from face offset were more strongly in the direction of the brow ($M = 0.4$, $SD = 0.153$) than the right eye ($M = 0.283$, $SD = 0.144$), $t(26) = 8.26$, $p < .001$ (adjusted $\alpha = .0033$), $g = 1.54$, 95% CI [1.09, 2.03], and right cheek ($M = 0.266$, $SD = 0.096$), $t(26) = 5.22$, $p < .001$ (adjusted $\alpha = .0036$), $g = 0.98$, 95% CI [0.58, 1.38], and marginally more in the direction of the brow than of the left eye ($M = 0.334$, $SD = 0.13$), $t(26) = 3.05$, $p = .0052$ (adjusted $\alpha = .0038$), $g = 0.57$, 95% CI [0.16, 1.0]. There was also a trend for reflexive first saccades to be more strongly in the direction of the mouth ($M = 0.326$, $SD = 0.138$) than the right cheek, $t(26) = 2.83$, $p = .009$ (though $p >$ adjusted α of .0042), $g = 0.53$, 95% CI [0.19, 0.84].

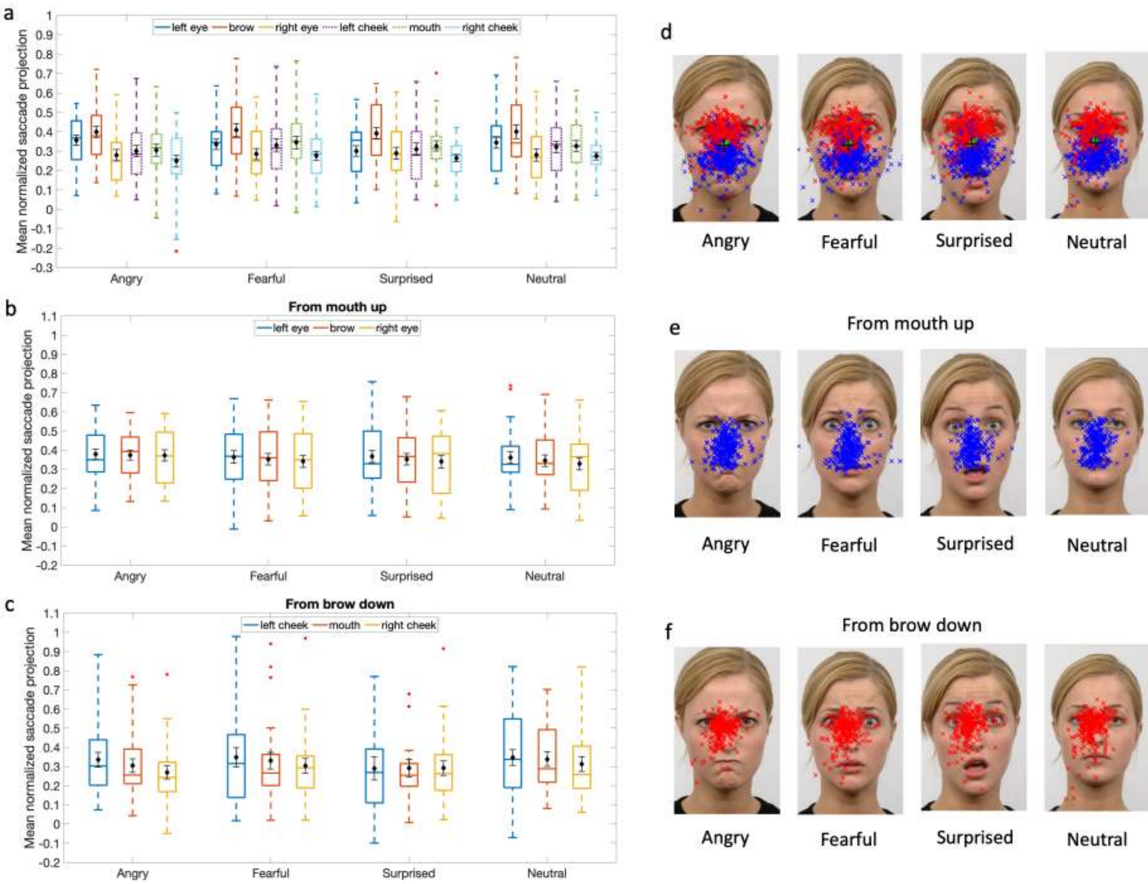

Figure 7. Mean normalized saccade projections as a function of the emotion expressed on the face and the target locations of interest (a – c) and saccade endpoints (d – f) for Experiment 2. The normalized saccade projection is a measure of the directional strength of reflexive first saccades (executed after face offset) towards target locations of interest, in this case (a) to 6 target locations, collapsed across initial fixation location (N = 27), (b) from the mouth center upwards towards the 3 upper-face target locations, and (c) from the central brow downwards towards the other 3 lower-face target locations. On each box, the central horizontal line indicates the median value across participants and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers. (Outliers, indicated with a '+', were defined as individual participant mean values more than 2 times the interquartile range away from the top or bottom of the box, and were not excluded from the data analyses.) Overlaid on the boxes are the mean values across participants, indicated by black diamonds, with error bars indicating the SEMs. (d) The endpoints for first saccades for each emotion, collapsed across participants and initial fixation location, plotted on example faces (which were not visible at the time of saccade execution). Blue crosses indicate endpoints of saccades from initial fixation on one of the lower-face features (mouth, cheeks) and red crosses indicate endpoints of saccades from initial fixation on one of the upper-face features (eyes, brow). The small green '+' indicate individual subject mean coordinates for their saccade endpoints, whereas the larger black '+' indicate the mean coordinates for first saccade endpoints, averaged over all trials for all subjects. The endpoints for first saccades upwards from fixation on the center of the mouth and downwards from the central brow are plotted separately in (e) and (f). The example face images in (d)-(f) are from the Langner et al. (2010) database. See the online article for the color version of this figure.

Next, as with Experiment 1, we tested whether the upward-directed saccades from the mouth more strongly targeted one or other (or both) of the eyes or the brow and whether this varied as a function of the emotional expression. The relevant normalized saccade projection data are summarized in Figure 7b. There were no reliable differences in the directional strength of reflexive saccades from the mouth towards the left or right eye or brow, either as a main effect of target location, $F(1.0, 26.03) = 1.17$, $p = .29$, $\eta_p^2 = .043$, $\eta_G^2 = .003$, or as a Target Location × Emotion interaction, $F(2.76, 71.83) = 1.55$, $p = .21$, $\eta_p^2 = .056$, $\eta_G^2 = .001$. There was also little evidence of a linear relationship in the normalized saccade projection values across the 3 target locations, $t(26) = -1.53$, $p = .133$ (polynomial contrast), as we had found in Experiment 1.

Finally, we tested whether the downward-directed saccades from the brow more strongly targeted the mouth or one or other (or both) of the cheeks and whether this varied as a function of emotion. The relevant normalized saccade projection data are summarized in Figure 7c. Similar to the previous analysis, there were no reliable differences in the directional strength of reflexive saccades from the brow towards the mouth or cheeks, either as a main effect of target location, $F(1.02, 21.45) = 1.45$, $p = .24$, $\eta_p^2 = .064$, $\eta_G^2 = .007$, or as a Target Location × Emotion interaction, $F(2.41, 50.56) = 1.19$, $p = .32$, $\eta_p^2 = .054$, $\eta_G^2 = .002$. There was also little evidence of a linear relationship in the normalized saccade projection values across the 3 target locations, $t(26) = -1.69$, $p = .1$ (polynomial contrast), collapsed across emotions, as we had found in Experiment 1. Nonetheless, Figure 7c suggests such a linear relationship is evident for angry faces and less so fearful faces, which was confirmed by polynomial contrasts: anger, $t(26) = -3.74$, $p < .001$; fear, $t(26) = -1.93$, $p = .059$.

Discussion

We replicated the finding from Experiment 1 that observers are better at discriminating angry expressions when the brow is projected to their fovea than when other facial features are projected foveally (and thus the brow is projected extrafoveally). This improvement was evidenced by greater accuracy, associated with a reduction in the number of misclassifications of angry faces as neutral, and shorter RTs. Importantly, a novel finding of Experiment 2 was that observers were better at discriminating fearful and surprised expressions when the mouth was projected to their fovea than when the brow or an eye or cheek was. This improved ability to classify fearful and surprised expressions when the mouth was at fixation was evidenced by greater accuracy and, for fear only and to a lesser extent, shorter RTs. The improved accuracy was principally associated with a reduction in the number of confusions between these two emotions. These results are consistent with previous work showing that the mouth distinguishes fearful from surprised as well as from neutral and angry expressions, whereas the eyes and brow do not distinguish between prototypical fearful and surprised expressions (Du, et al., 2014; Roy-Charland, et al., 2014; Smith, et al., 2005). Unlike in Experiment 1, we did not observe any variation in classification accuracy for neutral faces as a function of fixation location.

Although we found that proportionately more reflexive first saccades were made upwards from the mouth than downwards from the brow or eyes, replicating our finding in Experiment 1 and the basic pattern of results in previous studies (Boll & Gamer, 2014; Gamer & Büchel, 2009; Gamer, et al., 2010; Kliemann, et al., 2012; Kliemann, et al., 2010; Scheller, et al., 2012), we did not replicate the finding that this effect varies as a function of the expressed emotion. Moreover, as with Experiment 1, our saccade projection analyses did not indicate that those saccades preferentially target emotion-distinguishing facial features. In this second experiment, first saccades from face offset were, regardless of starting location and emotion, more strongly in the direction of the brow than the right eye and right cheek and marginally more strongly in the direction of the brow than of the left eye and in the direction of the mouth than of the right cheek. Compare this pattern of results with those of Experiment 1, which showed a clearer trend for first saccades to be more strongly in the direction of both the central face features (brow and mouth) than features on the right of the face (right eye and cheek). Moreover, the first saccades upwards from the center of the mouth did not exhibit the leftward directional bias observed in Experiment 1, that is, directed more strongly towards the central brow and even more strongly to the left eye, than to the right eye. In Experiment 1, first saccades downwards from the brow also showed this leftwards bias (for lower-face features) for all emotions, yet in Experiment 2 this was the case only for angry faces and, less clearly, for fearful faces.

General Discussion

We compared foveal and extrafoveal processing of features within a face by presenting face images to observers in a fixation-contingent manner for a time insufficient for a saccade, thus restricting the visual information available at fixation to a few key locations on the face. We had two main hypotheses: (1) Forcing a single fixation on a feature of a whole face that carries emotion-distinguishing content in medium-to-high spatial frequencies will enhance emotion recognition performance compared to having that feature in extrafoveal vision as the result of forcing fixation on a different, less informative feature. (2) When required to identify the expressed emotion, observers will preferentially saccade toward emotion-distinguishing features, especially those for which higher spatial frequency information would be most informative.

Our first hypothesis was clearly supported in the case of angry expressions: Across two experiments, observers were more accurate and faster at discriminating angry expressions when the central brow was projected to their fovea than when other facial features were projected foveally (and thus when the brow was projected extrafoveally). This improved accuracy was principally associated with a reduction in the number of misclassifications of angry faces as neutral. This novel finding is consistent with previous work showing the importance of high spatial-frequency information from the central brow in allowing observers to distinguish angry expressions from expressions of other basic emotions (Smith, et al., 2005). Performance in classifying fear and happiness was not influenced by whether their emotion-distinguishing features (eyes and mouth, respectively) were projected foveally or extrafoveally.

The main novel finding of Experiment 2 was that observers were much better able to distinguish between fearful and surprised expressions when the mouth of either expression was projected to their fovea than when an eye, cheek or mouth was. This improved performance was evidenced by fewer misclassifications of fear as surprise and of surprise as fear and, for fear only and to a lesser extent, shorter RTs. These results are consistent with our first hypothesis and with previous work showing that the mouth principally distinguishes fearful from surprised as well as from neutral and angry expressions, whereas the eyes and brow do not distinguish between prototypical fearful and surprised expressions (Du, et al., 2014; Roy-Charland, et al., 2014; Smith, et al., 2005). Certainly, in the face set we used, the shape of fearful and surprised mouths is quite different, and more of the teeth tend to be visible in fearful than in surprised expressions.

Why did enforcing fixation on emotion-distinguishing facial features enhance emotion classification performance for some emotions but not others? For expressions of happiness, ceiling effects are probably at play; that is, the very high accuracy and speed at which observers judged happy faces – a common finding in facial emotion classification studies (e.g., Calder, Young, et al., 2000; Calvo & Beltrán, 2014; Hugenberg & Sczesny, 2006; Kirita & Endo, 1995) – most probably did not leave enough space for the initial fixation location (e.g., mouth vs. brow) to have any effect. Ceiling effects are unlikely to be the whole story for happy faces, however. The happy face advantage is evident even when whole faces are presented in the extrafoveal visual field, upright or even inverted; indeed, the ability to classify happy faces remains relatively unaffected by extrafoveal presentation of those faces compared to expressions of other basic emotions and these effects can be attributed to the extrafoveal processing of the visually salient and diagnostic smiling mouth (Calvo, Fernández-Martín, & Nummenmaa, 2014; Calvo, Nummenmaa, & Avero, 2010). We suggest that, similar to when whole faces appear extrafoveally, the extrafoveal processing of the mouth region within a centrally presented whole face, is sufficient for the extraction of the diagnostic smile indicating happiness.

Ceiling effects likely played less of a role in the results obtained for fearful expressions in Experiment 1 and for fearful and surprised expressions in Experiment 2, given the observed level of performance in those experiments. In Experiment 1, enforced fixation location did have some effect on classification performance for fearful faces, but fixation on the emotion-distinguishing eyes did not improve accuracy compared to any of the other fixation locations and reduced RTs only relative to fixation on the brow. Observers were more accurate when fixating the mouth than the brow of fearful faces, and this small improvement was principally associated with a reduction in the number of misclassifications of fearful faces as angry. Thus, rather than fixation on an eye specifically enhancing classification of fearful expressions, it was more that fixation on the brow degraded classification of fearful expressions relative to fixation on an eye or the mouth, both of which are regions that observers use to classify fearful expressions, depending on the particular comparison emotions employed: In free-viewing tasks, observers rely on the mid-to-high spatial frequency information in the eyes for the successful classification of fearful faces across a range of comparisons with other emotions, but also rely on mid-to-low spatial frequency information in the mouth when distinguishing fearful from neutral and angry faces, and around the corners of the mouth when distinguishing fearful from happy and neutral faces (Smith & Merlusca, 2014). Extrafoveal processing of this mid-to-low spatial frequency information, in combination with the little or no added benefit of foveal vs. extrafoveal processing of low spatial frequency information, might account for why we did not observe a relative benefit of forced fixation on the eyes of fearful faces. In Experiment 2, fixation on an eye of fearful faces did not enhance emotion classification performance compared to fixation on the other tested locations (brow, mouth, cheeks), most probably for the same reason as we indicated for Experiment 1, but also because Experiment 2 included expressions of surprise and enforcing fixation on the mouth allowed observers to more readily distinguish between expressions of these two emotions.

Our hypothesis relating to emotion classification performance is based on the assumption that the distance between a feature and the fixation location is directly related to the spatial frequencies that can be extracted by the visual system. Visual processing and performance on a variety of visual tasks do not vary uniformly between the lower and upper and between the left and right visual fields, however. Performance on some visual tasks tends to be better at equivalent eccentricities along the horizontal than vertical meridian (e.g., Carrasco, Evert, Chang, & Katz, 1995; Carrasco, Talgar, & Cameron, 2001; Rijsdijk, Kroon, & Vanderwildt, 1980; Rovamo & Virsu, 1979). There is also a lower visual field advantage across a range of psychophysical tasks, such as those measuring contrast sensitivities, visual acuity, spatial resolution, and the effects of visual crowding, which is evident even at the same eccentricities and becomes more pronounced with increased eccentricity and spatial frequency, but is mostly or only evident along the vertical meridian (e.g., Carrasco, et al., 2001; Edgar & Smith, 1990; S. He, Cavanagh, & Intriligator, 1996; Liu, Heeger, & Carrasco, 2006; Talgar & Carrasco, 2002).

A lower visual field advantage in processing spatial frequencies, especially around the vertical meridian, could help explain some of our results. Fixation on one of the upper-face features puts the lower face in the lower visual field, whereas fixation on one of the lower-face features puts the upper-face in the upper visual field. Now, consider the case of angry faces in our experiments. When fixation is enforced at the central brow, the center of the mouth is directly below, whereas with fixation at the center of the mouth, the central brow is directly above, at an eccentricity of 5.6^{o}. At this eccentricity, the lower visual field advantage (in an orientation discrimination task) becomes evident for spatial frequencies at around 10 cycles per degree (Carrasco, et al., 2001). Thus, with fixation at the brow, observers might be able to extract more emotion-relevant mid-to-high spatial frequency information from the extrafoveal mouth than they can from the extrafoveal brow with fixation at the mouth. A recognition advantage for fixating the angry brow therefore might not be due entirely to the foveal advantage for processing high spatial frequencies at the brow. A similar argument does not apply so readily to our emotion classification findings for fear and surprise in Experiment 2, however. Even though the mouth distinguishes between fearful and surprised expressions, enforcing fixation at the brow and thus putting the mouth on the lower vertical meridian did not lead to better emotion classification performance relative to fixation of an eye or cheek (see Figure 5). More generally, a lower visual field advantage for processing task-relevant spatial frequencies is in tension with evidence showing that, for some tasks, including a variety of face perception tasks, there is an upper visual field advantage (e.g., Carlei, Framorando, Burra, & Kerzel, 2017; Hagenbeek & Van Strien, 2002; Quek & Finkbeiner, 2014a; Quek & Finkbeiner, 2014b, 2016), which, for faces and some objects, might be related to their typical visual-field positioning in everyday environments (Kaiser & Cichy, 2018; Kaiser, Quek, Cichy, & Peelen, 2019). Indeed, even isolated face parts (eyes, mouth) are better recognized and more strongly represented in right inferior occipital gyrus when they were presented at typical, rather than reversed, visual field locations (de Haas et al., 2016). There are also left visual-field advantages for many aspects of face perception, including emotion perception (e.g., Adolphs, 2002; Burt & Perrett, 1997) and recent evidence indicates that low spatial frequencies are not generally sufficient to induce a left visual field bias in emotion perception under free-viewing conditions (Hausmann, Innes, Birch, & Kentridge, 2019). These issues deserve further investigation.

Our study also addressed the hypothesis that, when required to identify the expressed emotion, observers' initial eye movements will reflect the location of task-relevant diagnostic features – specifically, they will preferentially saccade toward emotion-distinguishing facial features, especially those features for which higher spatial frequency information is most informative. The authors of some previous studies that have employed the brief-fixation paradigm have sought to address the preferential seeking out of emotion-distinguishing features with a measure of the relative proportion of saccades upward from initial fixation on the mouth ('toward the eyes') and downward from initial fixation on an eye or on the midpoint between the eyes ('toward the mouth'). In both of our experiments, we replicated a key finding of most of those previous studies, namely, more upward saccades from enforced fixation on the mouth than downward saccades from enforced fixation on or between the eyes (Boll & Gamer, 2014; Gamer & Büchel, 2009; Gamer, et al., 2013; Gamer, et al., 2010; Kliemann, et al., 2012; Kliemann, et al., 2010; Scheller, et al., 2012). In Experiment 1 but not in Experiment 2, we also replicated the finding that this difference in upward versus downward saccades varies depending on the displayed emotion, in our case being absent for happy faces, evident for fearful faces, and larger for angry and neutral faces.

A limitation of this proportion saccade measure, however, is that it is unable to provide a measure of the extent to which eye movements are directed toward specific facial features. This is compounded by the way in which this measure is typically used, whereby the eye and mouth regions to which saccades are allegedly directed are geographically large regions of the face, especially horizontally. In an attempt to provide a more spatially precise measure of the extent to which reflexive first saccades target diagnostic features, we calculated a saccade projection measure by projecting the vector of each first saccade on to the vectors from the enforced fixation location to each of the other facial locations of interest, normalized for the length of the target vector (given that the target locations vary in distance from a given fixation location). Using this measure, we did not find any support for the hypothesis that observers' first saccades from initial fixation on the face will seek out emotion-distinguishing features (e.g., the central brow for anger, the mouth for fear vs. surprise). It is possible that this is because the very short image duration – which was necessary for preventing additional fixations on the face – did not allow sufficient time for saccades to be planned with high spatial certainty and thus that the saccades that we analyzed only reflect a general directional tendency. Our saccade projection measure provided evidence of some such general directional tendencies.

Collapsed over starting location and emotion, first saccades were more strongly in the direction of the central face features (brow and mouth) than features on the right of the face (right eye and cheek). This finding is more consistent with previously reported 'center-of-gravity' effects, that is, a strong tendency for first saccades to be to the geometric center of scenes or configurations (e.g., Bindemann, et al., 2010; Findlay, 1982; P. Y. He & Kowler, 1989; Tatler, 2007), including faces (Bindemann, et al., 2009). Moreover, there was no evidence that this center-of-gravity effect in our data was modified by the emotion on the face (which is also illustrated in our plots of saccade endpoints, in Figures 4d and 7d). Nevertheless, there is evidence suggesting that simply averaging saccade endpoint data across start positions will tend to artificially regress those endpoint locations toward the center of the face (Arizpe, Kravitz, Yovel, & Baker, 2012), so it is important also to examine saccades for individual start positions.

When considering only upwards saccades from fixation on the mouth and downwards saccades from fixation on the brow, there was no evidence that those saccades preferentially targeted emotion-specific distinguishing features. In Experiment 1, those saccades showed a directional bias towards the central (brow or mouth) and left-sided (eye or check) facial features. In Experiment 2, by contrast, those saccades were more evenly directionally distributed between the central and lateral features, though the central/leftward bias was still evident for downward saccades from the brow for angry faces and, though less clearly, for fearful faces. This central/leftward bias for reflexive saccades from the center of the brow and mouth might reflect one or more of (a) the center of gravity effect, (b) the left visual field/right hemisphere advantage in emotion perception (e.g., Adolphs, 2002; Burt & Perrett, 1997; Butler et al., 2005), and (c) the strong tendency for first saccades onto a face to target a location below the eyes, just to the left of face center, which is also the optimal initial fixation point for determining a face's emotional expression (Peterson & Eckstein, 2012).

A possible limitation of our study, raised by one of the reviewers, is that unique images (e.g., person A with an angry expression) were viewed multiple times by a given participant over the course of the experiment and so emotion classification performance and perhaps even the direction of gaze shifts might be due in part to memory effects (i.e., as the result of the participant having remembered, implicitly or explicitly, a given image). We believe such a confound is unlikely to have affected our results. For, each image was presented for only 82 ms and repeated only 4 times, separated by an average of 95 other images, within a fairly rapid sequence of trials; moreover, the 4 repeated images were at different locations relative to fixation and most repetitions occurred between rather than within stimulus blocks, with each block being separated in time by at least several minutes. Nonetheless, to check the impact of memory or practice effects on emotion classification performance, we conducted exploratory analyses on the emotion classification accuracy data (see Supplementary Materials). Although accuracy tended to increase linearly with image repetition, this was only the case for angry faces in Experiment 1 and for angry and fearful faces in Experiment 2. Crucially, in both experiments the interaction between emotion and fixation location, which is important for testing our hypothesis that fixation on emotion-distinguishing facial features will enhance emotion classification performance, was not modified by a further interaction with image repetition. Although the results of these analyses have to be treated with caution, they suggest that image repetition was not a serious confound.

Another limitation of our study is that we tested a limited combination of emotions. Future research should test different combinations of emotions, incorporating those that have not yet been used widely or at all in the brief-fixation paradigm (e.g., disgust, sadness) and combinations that reflect common misclassifications (as in our use of fear vs. surprise). Our ongoing work is doing just that. Future work could also investigate the impact of methodological differences between our study and similar, previously published studies. Notably, in our study, colored images of faces that were not cropped to exclude extra-facial features were presented for 82 ms, whereas other studies typically presented greyscale, cropped faces for 150 ms. It is possible, for example, that the tendency of first saccades to be directed towards the center is more pronounced in uncropped images with more visible borders between the image and the background.

We suggest that the findings reported in this study reflect the interplay between two factors: first, that the integration of task-relevant information across the face is constrained by the varying spatial resolution of visual processing across the retina (Peterson & Eckstein, 2012), that is, constrained by the differences between foveal and extrafoveal processing (Atkinson & Smithson, 2013; Rosenholtz, 2016; Strasburger, et al., 2011); second, that observers make use of different visual information from expressive faces (different features and or different spatial frequencies) depending on task demands, particularly the specific emotions and response options that are pitted against each other (Smith & Merlusca, 2014). One avenue for future research, which we are currently following, is to investigate whether the differences in emotion classification accuracy between foveal and extrafoveal processing of emotion-distinguishing features is fully accounted for by extrafoveal blurring, or whether some other factor (such as crowding) also contributes.

References

Adolphs, R. (2002). Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews, 1*(1), 21-62. http://dx.doi.org/10.1177/1534582302001001003

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature, 433*(7021), 68-72. http://dx.doi.org/10.1038/nature03086

Arizpe, J., Kravitz, D. J., Yovel, G., & Baker, C. I. (2012). Start position strongly influences fixation patterns during face processing: Difficulties with eye movements as a measure of information use. *PLoS One, 7*(2), e31106. http://dx.doi.org/10.1371/journal.pone.0031106

Arizpe, J., Walsh, V., Yovel, G., & Baker, C. I. (2017). The categories, frequencies, and stability of idiosyncratic eye-movement patterns to faces. *Vision Research, 141*, 191-203. http://dx.doi.org/10.1016/j.visres.2016.10.013

Atkinson, A. P., & Smithson, H. E. (2013). Distinct contributions to facial emotion perception of foveated versus nonfoveated facial features. *Emotion Review, 5*(1), 30-35. http://dx.doi.org/10.1177/1754073912457226

Atkinson, A. P., & Smithson, H. E. (2019). Data for: The impact on emotion classification performance and gaze behavior of foveal vs. extrafoveal processing of facial features (1 ed.). https://osf.io/vduzt/.

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods, 37*(3), 379-384. http://dx.doi.org/10.3758/BF03192707

Beaudry, O., Roy-Charland, A., Perron, M., Cormier, I., & Tapp, R. (2013). Featural processing in recognition of emotional facial expressions. *Cognition & Emotion, 28*(3), 416-432. http://dx.doi.org/10.1080/02699931.2013.833500

Bindemann, M., Scheepers, C., & Burton, A. M. (2009). Viewpoint and center of gravity affect eye movements to human faces. *Journal of Vision, 9*(2), 1-16. http://dx.doi.org/10.1167/9.2.7

Bindemann, M., Scheepers, C., Ferguson, H. J., & Burton, A. M. (2010). Face, body, and center of gravity mediate person detection in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 36*(6), 1477-1485. http://dx.doi.org/10.1037/a0019057

Boll, S., & Gamer, M. (2014). 5-HTTLPR modulates the recognition accuracy and exploration of emotional facial expressions. *Frontiers in Behavioral Neuroscience, 8*. http://dx.doi.org/10.3389/fnbeh.2014.00255

Boucher, J. D., & Ekman, P. (1975). Facial areas and emotional information. *Journal of Communication, 25*(2), 21-29.

Burt, D. M., & Perrett, D. I. (1997). Perceptual asymmetries in judgements of facial attractiveness, age, gender, speech and expression. *Neuropsychologia, 35*(5), 685-693. http://dx.doi.org/10.1016/S0028-3932(96)00111-X

Butler, S., Gilchrist, I. D., Burt, D. M., Perrett, D. I., Jones, E., & Harvey, M. (2005). Are the perceptual biases found in chimeric face processing reflected in eye-movement patterns? *Neuropsychologia, 43*(1), 52-59. http://dx.doi.org/10.1016/j.neuropsychologia.2004.06.005

Calder, A. J., Rowland, D., Young, A. W., Nimmo-Smith, I., Keane, J., & Perrett, D. I. (2000). Caricaturing facial expressions. *Cognition, 76*(2), 105-146. http://dx.doi.org/10.1016/S0010-0277(00)00074-3

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*(2), 527-551. http://dx.doi.org/10.I037/TO96-1523.26.2.527

Calvo, M. G., & Beltrán, D. (2014). Brain lateralization of holistic versus analytic processing of emotional facial expressions. *NeuroImage, 92*(0), 237-247. http://dx.doi.org/10.1016/j.neuroimage.2014.01.048

Calvo, M. G., Fernández-Martín, A., & Nummenmaa, L. (2014). Facial expression recognition in peripheral versus central vision: role of the eyes and the mouth. *Psychological Research, 78*(2), 180-195. http://dx.doi.org/10.1007/s00426-013-0492-x

Calvo, M. G., Nummenmaa, L., & Avero, P. (2010). Recognition advantage of happy faces in extrafoveal vision: Featural and affective processing. *Visual Cognition, 18*(9), 1274 - 1297. http://dx.doi.org/10.1080/13506285.2010.481867

Carlei, C., Framorando, D., Burra, N., & Kerzel, D. (2017). Face processing is enhanced in the left and upper visual hemi-fields. *Visual Cognition, 25*(7-8), 749-761. http://dx.doi.org/10.1080/13506285.2017.1327466

Carrasco, M., Evert, D. L., Chang, I., & Katz, S. M. (1995). The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Perception & Psychophysics, 57*(8), 1241-1261. http://dx.doi.org/10.3758/Bf03208380

Carrasco, M., Talgar, C. P., & Cameron, E. L. (2001). Characterizing visual performance fields: effects of transient covert attention, spatial frequency, eccentricity, task and set size. *Spatial Vision, 15*(1), 61-75. http://dx.doi.org/10.1163/15685680152692015

Cumming, G., & Calin-Jageman, R. (2016). *Introduction to the New Statistics: Estimation, Open Science, and Beyond*. New York, NY: Routledge.

de Haas, B., Schwarzkopf, D. S., Alvarez, I., Lawson, R. P., Henriksson, L., Kriegeskorte, N., et al. (2016). Perception and processing of faces in the human brain is tuned to typical feature locations. *The Journal of Neuroscience, 36*(36), 9289-9302. http://dx.doi.org/10.1523/jneurosci.4131-14.2016

Du, S. C., Tao, Y., & Martinez, A. M. (2014). Compound facial expressions of emotion. *Proceedings of the National Academy of Sciences of the United States of America, 111*(15), E1454-E1462. http://dx.doi.org/10.1073/pnas.1322355111

Edgar, G. K., & Smith, A. T. (1990). Hemifield differences in perceived spatial frequency. *Perception, 19*(6), 759-766. http://dx.doi.org/10.1068/p190759

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

Fang, F., Clapham, P. J., & Chung, K. C. (2011). A systematic review of interethnic variability in facial dimensions. *Plastic and Reconstructive Surgery, 127*(2), 874-881. http://dx.doi.org/10.1097/PRS.0b013e318200afdb

Findlay, J. M. (1982). Global visual processing for saccadic eye-movements. *Vision Research, 22*(8), 1033-1045. http://dx.doi.org/10.1016/0042-6989(82)90040-2

Fuchs, A. F., Kaneko, C. R., & Scudder, C. A. (1985). Brainstem control of saccadic eye movements. *Annual Review of Neuroscience, 8*, 307-337. http://dx.doi.org/10.1146/annurev.ne.08.030185.001515

Gamer, M., & Büchel, C. (2009). Amygdala activation predicts gaze toward fearful eyes. *The Journal of Neuroscience, 29*(28), 9123-9126. http://dx.doi.org/10.1523/jneurosci.1883-09.2009

Gamer, M., Schmitz, A. K., Tittgemeyer, M., & Schilbach, L. (2013). The human amygdala drives reflexive orienting towards facial features. *Current Biology, 23*(20), R917-R918. http://dx.doi.org/http://dx.doi.org/10.1016/j.cub.2013.09.008

Gamer, M., Zurowski, B., & Büchel, C. (2010). Different amygdala subregions mediate valence-related and attentional effects of oxytocin in humans. *Proceedings of the National Academy of Sciences, 107*(20), 9400-9405. http://dx.doi.org/10.1073/pnas.1000985107

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin, 75*(6), 424-429. http://dx.doi.org/10.1037/H0031246

Hagenbeek, R. E., & Van Strien, J. W. (2002). Left–right and upper–lower visual field asymmetries for face matching, letter naming, and lexical decision. *Brain and Cognition, 49*(1), 34-44. http://dx.doi.org/10.1006/brcg.2001.1481

Hall, E. T. (1966). *The hidden dimension*. Garden City, NY: Doubleday.

Hausmann, M., Innes, B. R., Birch, Y. K., & Kentridge, R. W. (2019). Laterality and (in)visibility in emotional face perception: Manipulations in spatial frequency content. *Emotion*. http://dx.doi.org/10.1037/emo0000648

He, P. Y., & Kowler, E. (1989). The role of location probability in the programming of saccades: Implications for center-of-Gravity tendencies. *Vision Research, 29*(9), 1165-1181. http://dx.doi.org/10.1016/0042-6989(89)90063-1

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383*(6598), 334-337. http://dx.doi.org/10.1038/383334a0

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition, 33*(1), 98-106. http://dx.doi.org/10.3758/BF03195300

Hsiao, J. H., & Cottrell, G. W. (2008). Two fixations suffice in face recognition. *Psychological Science, 19*(10), 998-1006. http://dx.doi.org/10.1111/j.1467-9280.2008.02191.x

Hugenberg, K., & Sczesny, S. (2006). On wonderful women and seeing smiles: Social categorization moderates the happy face response latency advantage. *Social Cognition, 24*(5), 516-539. http://dx.doi.org/10.1521/soco.2006.24.5.516

Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology, 19*(18), 1543-1548. http://dx.doi.org/10.1016/j.cub.2009.07.051

Kaiser, D., & Cichy, R. M. (2018). Typical visual-field locations facilitate access to awareness for everyday objects. *Cognition, 180*, 118-122. http://dx.doi.org/10.1016/j.cognition.2018.07.009

Kaiser, D., Quek, G. L., Cichy, R. M., & Peelen, M. V. (2019). Object vision in a structured world. *Trends in Cognitive Sciences, 23*(8), 672-685. http://dx.doi.org/10.1016/j.tics.2019.04.013

Kirby, K. N., & Gerlanc, D. (2013). BootES: An R package for bootstrap confidence intervals on effect sizes. *Behavior Research Methods, 45*(4), 905-927. http://dx.doi.org/10.3758/s13428-013-0330-5

Kirita, T., & Endo, M. (1995). Happy face advantage in recognizing facial expressions. *Acta Psychologica, 89*(2), 149-163. http://dx.doi.org/10.1016/0001-6918(94)00021-8

Kliemann, D., Dziobek, I., Hatri, A., Baudewig, J., & Heekeren, H. R. (2012). The role of the amygdala in atypical gaze on emotional faces in Autism Spectrum Disorders. *The Journal of Neuroscience, 32*(28), 9469-9476. http://dx.doi.org/10.1523/jneurosci.5294-11.2012

Kliemann, D., Dziobek, I., Hatri, A., Steimke, R., & Heekeren, H. R. (2010). Atypical reflexive gaze patterns on emotional faces in autism spectrum disorders. *The Journal of Neuroscience, 30*(37), 12281-12287. http://dx.doi.org/10.1523/jneurosci.0688-10.2010

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*. http://dx.doi.org/10.3389/fpsyg.2013.00863

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion, 24*(8), 1377-1388. http://dx.doi.org/10.1080/02699930903485076

Liu, T., Heeger, D. J., & Carrasco, M. (2006). Neural correlates of the visual vertical meridian asymmetry. *Journal of Vision, 6*(11), 1294-1306. http://dx.doi.org/10.1167/6.11.12

Liversedge, S. P., Rayner, K., White, S. J., Vergilino-Perez, D., Findlay, J. M., & Kentridge, R. W. (2004). Eye movements when reading disappearing text: is there a gap effect in reading? *Vision Research, 44*(10), 1013-1024. http://dx.doi.org/10.1016/J.Visres.2003.12.002

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

McKone, E. (2009). Holistic processing for faces operates over a wide range of sizes but is strongest at identification rather than conversational distances. *Vision Research, 49*(2), 268-283. http://dx.doi.org/10.1016/j.visres.2008.10.020

Mehoudar, E., Arizpe, J., Baker, C. I., & Yovel, G. (2014). Faces in the eye of the beholder: Unique and stable eye scanning patterns of individual observers. *Journal of Vision, 14*(7). http://dx.doi.org/10.1167/14.7.6

Neath, K. N., & Itier, R. J. (2014). Facial expression discrimination varies with presentation time but not with fixation on features: A backward masking study using eye-tracking. *Cognition & Emotion, 28*(1), 115-131. http://dx.doi.org/10.1080/02699931.2013.812557

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods, 42*(1), 188-204. http://dx.doi.org/10.3758/brm.42.1.188

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods, 8*(4), 434-447. http://dx.doi.org/10.1037/1082-989x.8.4.434

Oruç, İ., & Barton, J. J. S. (2010). Critical frequencies in the perception of letters, faces, and novel shapes: Evidence for limited scale invariance for faces. *Journal of Vision, 10*(12). http://dx.doi.org/10.1167/10.12.20

Peli, E., Yang, J., & Goldstein, R. B. (1991). Image invariance with changes in size: the role of peripheral contrast thresholds. *Journal of the Optical Society of America A, 8*(11), 1762-1774. http://dx.doi.org/10.1364/JOSAA.8.001762

Peterson, M. F., & Eckstein, M. P. (2012). Looking just below the eyes is optimal across face recognition tasks. *Proceedings of the National Academy of Sciences, 109*(48), E3314-E3323. http://dx.doi.org/10.1073/pnas.1214269109

Peterson, M. F., & Eckstein, M. P. (2013). Individual differences in eye movements during face identification reflect observer-specific optimal points of fixation. *Psychological Science, 24*(7), 1216-1225. http://dx.doi.org/10.1177/0956797612471684

Quek, G. L., & Finkbeiner, M. (2014a). Face-perception is superior in the upper visual field: Evidence from masked priming. *Visual Cognition, 22*(8), 1038-1042. http://dx.doi.org/10.1080/13506285.2014.960732

Quek, G. L., & Finkbeiner, M. (2014b). Face-sex categorization is better above fixation than below: Evidence from the reach-to-touch paradigm. *Cognitive Affective & Behavioral Neuroscience, 14*(4), 1407-1419. http://dx.doi.org/10.3758/s13415-014-0282-y

Quek, G. L., & Finkbeiner, M. (2016). The upper-hemifield advantage for masked face processing: Not just an attentional bias. *Attention, Perception, & Psychophysics, 78*(1), 52-68. http://dx.doi.org/10.3758/s13414-015-0965-7

Rijsdijk, J. P., Kroon, J. N., & Vanderwildt, G. J. (1980). Contrast sensitivity as a function of position on the retina. *Vision Research, 20*(3), 235-241. http://dx.doi.org/10.1016/0042-6989(80)90108-X

Robson, J. G., & Graham, N. (1981). Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Research, 21*(3), 409-418. http://dx.doi.org/10.1016/0042-6989(81)90169-3

Rosenholtz, R. (2016). Capabilities and limitations of peripheral vision. *Annual Review of Vision Science, 2*(1), 437-457. http://dx.doi.org/10.1146/annurev-vision-082114-035733

Ross, D. A., & Gauthier, I. (2015). Holistic processing in the composite task depends on face size. *Visual Cognition, 23*(5), 533-545. http://dx.doi.org/10.1080/13506285.2015.1049678

Rovamo, J., & Virsu, V. (1979). An estimation and application of the human cortical magnification factor. *Experimental Brain Research, 37*(3), 495-510. http://dx.doi.org/10.1007/bf00236819

Roy-Charland, A., Perron, M., Beaudry, O., & Eady, K. (2014). Confusion of fear and surprise: A test of the perceptual-attentional limitation hypothesis with eye movement monitoring. *Cognition and Emotion, 28*(7), 1214-1222. http://dx.doi.org/10.1080/02699931.2013.878687

Saslow, M. G. (1967). Effects of components of displacement-step stimuli upon latency for saccadic eye movement. *Journal of the Optical Society of America, 57*(8), 1024-1029.

Scheller, E., Büchel, C., & Gamer, M. (2012). Diagnostic features of emotional expressions are processed preferentially. *PLoS One, 7*(7), e41792. http://dx.doi.org/10.1371/journal.pone.0041792

Schurgin, M. W., Nelson, J., Iida, S., Ohira, H., Chiao, J. Y., & Franconeri, S. L. (2014). Eye movements during emotion recognition in faces. *Journal of Vision, 14*(13). http://dx.doi.org/10.1167/14.13.14

Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science, 16*(3), 184-189. http://dx.doi.org/10.1111/j.0956-7976.2005.00801.x

Smith, M. L., & Merlusca, C. (2014). How task shapes the use of information during facial expression categorizations. *Emotion, 14*(3), 478-487. http://dx.doi.org/10.1037/a0035588

Sperling, G. (1965). Temporal and spatial visual masking. I. Masking by impulse flashes. *Journal of the Optical Society of America, 55*(5), 541-&. http://dx.doi.org/10.1364/Josa.55.000541

Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision, 11*(5). http://dx.doi.org/10.1167/11.5.13

Talgar, C. P., & Carrasco, M. (2002). Vertical meridian asymmetry in spatial resolution: Visual and attentional factors. *Psychonomic Bulletin & Review, 9*(4), 714-722. http://dx.doi.org/10.3758/Bf03196326

Tanaka, J. W., Kaiser, M. D., Butler, S., & Le Grand, R. (2012). Mixed emotions: Holistic and analytic perception of facial expressions. *Cognition & Emotion, 26*(6), 961-977. http://dx.doi.org/10.1080/02699931.2011.630933

Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision, 7*(14). http://dx.doi.org/10.1167/7.14.4

Vaidya, A. R., Jin, C., & Fellows, L. K. (2014). Eye spy: The predictive value of fixation patterns in detecting subtle and extreme emotions from faces. *Cognition, 133*(2), 443-456. http://dx.doi.org/10.1016/j.cognition.2014.07.004

Van Belle, G., De Graef, P., Verfaillie, K., Rossion, B., & Lefèvre, P. (2010). Face inversion impairs holistic perception: Evidence from gaze-contingent stimulation. *Journal of Vision, 10*(5), 1-13. http://dx.doi.org/10.1167/10.5.10

Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior, 17*(1), 3-28. http://dx.doi.org/10.1007/BF00987006

Walker, R., Kentridge, R. W., & Findlay, J. M. (1995). Independent contributions of the orienting of attention, fixation offset and bilateral stimulation on human saccadic latencies. *Experimental Brain Research, 103*(2), 294-310. http://dx.doi.org/10.1007/BF00231716

Wandell, B. A. (1995). *Foundations of vision*. Sunderland, MA: Sinauer Associates.

White, M. (2000). Parts and wholes in expression recognition. *Cognition and Emotion, 14*(1), 39-60. http://dx.doi.org/10.1080/026999300378987

Woodworth, R. S. (1938). *Experimental psychology*. New York: Henry Holt.

Wurtz, R. H. (1996). Vision for the control of movement. The Friedenwald Lecture. *Investigative Ophthalmology & Visual Science, 37*(11), 2130-2145.

Yang, N., Shafai, F., & Oruc, I. (2014). Size determines whether specialized expert processes are engaged for recognition of faces. *Journal of Vision, 14*(8). http://dx.doi.org/10.1167/14.8.17

Yarbus, A. L. (1967). *Eye movements and vision* (B. Haigh, Trans.). New York: Plenum Press.

Footnotes

1. Although 'fovea' is sometimes used to refer to the larger region of the retina encompassing the central 5.2° of the visual field, it is also sometimes used to refer to the smaller area (the 'central fovea') that contains the highest concentration of cone cells but no rod cells and which encompasses the central 1.7° of the visual field (Wandell, 1995). We here follow this latter convention. Importantly, for our purposes, visual acuity and contrast sensitivity decline and crowding increases even within the central 5.2° of the visual field (i.e., with < 2.6° eccentricity) (Peli, Yang, & Goldstein, 1991; Robson & Graham, 1981; Rosenholtz, 2016).

2. In light of the strong contribution of configural or holistic processing to facial emotion perception (e.g., Calder, Young, et al., 2000; Calvo & Beltrán, 2014; Tanaka, et al., 2012; White, 2000), we prefer to refer to certain facial features as being 'emotion-distinguishing' or 'relatively more (or less) informative' than as being 'diagnostic' (cf. e.g., Smith, et al., 2005).

3. It should be noted that Neath and Itier's (2014) findings are not directly comparable to those of the other studies, for two reasons. First, the target expressions in Neath and Itier's experiments were immediately followed by either an upright or inverted neutral face mask (for 150 ms), which was not the case in the other studies reviewed here. Backward pattern masking is commonly used in studies involving brief stimulus presentations; since it is argued that without the mask (Sperling, 1965) the target stimulus will remain available to the viewer in iconic memory (Neisser, 1967). Second, Neath and Itier did not present proportion correct data but instead assessed accuracy using A', a nonparametric analogue of the d' sensitivity measure, which takes into account false alarm rates as well as hit rates, thus providing a measure of discrimination performance independent of any response biases (e.g., Grier, 1971; Macmillan &

Creelman, 2005). Moreover, there are two further details of Neath and Itier's (2014) study that could have contributed to their pattern of results. First, the faces were always presented at the same location on the screen, with the immediately preceding fixation cross always appearing in approximately the same screen location for a given target facial feature. Thus, participants were effectively cued in advance to what feature they would be fixating. Second, the fixation location on the forehead was midway between the edge of the hair line and the nasion, which is higher up the forehead than the central brow region that Smith et al.'s (2005) work using the Bubbles technique identified as diagnostic of angry and sad expressions, which Neath and Itier did not use in their study anyway.

4. The "unbiased hit rate" (H_u) accounts for response biases in classification experiments with multiple response options (Wagner, 1993). H_u for each participant is calculated as the squared frequency of correct responses for a target emotion in a particular condition divided by the product of the number of stimuli in that condition representing this emotion and the overall frequency that that emotion category is chosen for that condition. H_u ranges from 0 to 1, with 1 indicating that all stimuli in a given condition representing a particular emotion have been correctly identified and that that emotion label has never been falsely selected for a different emotion.

5. We did not also analyze the normalized saccade projection values for the downward-directed saccades from each eye, for two reasons. First, there were considerably fewer such saccades in these conditions, especially given that each eye had only half the number of trials as the mouth or brow. Second, the brow is more comparable to the mouth as a saccade start location, given that it is directly above the mouth location (see Figure 1) and is closer to the start location between the

eyes used by most other relevant previous studies (Boll & Gamer, 2014; Gamer, et al., 2010; Kliemann, et al., 2012; Kliemann, et al., 2010; Scheller, et al., 2012) than is either eye.