Eye-tracking the own-gender bias in face recognition: Other-gender faces are viewed differently to own-gender faces

Tsoey Wun Man[1] & Peter J Hills[2]

1 - Department of Psychology, Anglia Ruskin University, Cambridge, UK, CB1 1PT. Email: tsoey.man@student.anglia.ac.uk

2 - Address for correspondence: Department of Psychology, Bournemouth University, Talbot Campus, Fern Barrow, Poole, Dorset, UK, BH12 5BB. Email: phills@bournemouth.ac.uk. Tel: 01202 961570

Word count: 5873

Abstract

Research on the own-gender bias in face recognition has indicated an asymmetrical effect: an effect found only in women (Lovén, Herlitz, & Rehnman, 2011). We investigated the own-gender bias, using an eye-tracker to examine whether the own-gender bias is associated with differential processing strategies. We found an own-gender bias in our female participants, consistent with Herlitz and Lovén (2013). Our eye-tracking analysis indicated different scanning behaviours when processing own- and other-gender faces, with longer and more fixations to the eyes when viewing own-gender faces. Our results favour the socio-cognitive model, whilst acknowledging the role of perceptual expertise in the own-gender bias.

Keywords: Face recognition; own-gender bias; eye-tracking; own-group biases; gender effects

Eye-tracking the own-gender bias in face recognition

Faces play important roles in everyday social interactions as they provide crucial information about people, such as identity and their emotional state (Kanwisher & Yovel, 2006). Coding this information is relatively quick given the expertise that humans have with face processing. Nevertheless, research has indicated that there are strong own-group biases; specifically age (Rhodes & Anastasi, 2012), and ethnicity (Feingold, 1914; O'Toole, Deffenbacher, Valentin, & Abdi, 1994; Meissner & Brigham, 2001). The own-group biases are where individuals are significantly quicker and more accurate at recognising faces of their own group relative to out-group faces. These biases are typically robust, though the own-ethnicity bias is sometimes found to be stronger in White participants than Asian participants (Valentine & Endo, 1992). However, the own-gender bias in face recognition is less robust than the other biases. Studies have indicated that this bias is asymmetrical: an own-gender bias has been found in women but not in men; i.e. women were significantly more accurate at recognising female faces than male faces but men show more equal recognition for male and female faces (e.g., Lovén, Herlitz, & Rehnman, 2011; Rehnman & Herlitz, 2007; see also Rehnman & Herlitz, 2006). The own-gender bias appears to be present in women regardless of the difficulty of the experimental task or how much of a face is presented (Lewin & Herlitz, 2002; Lovén et al., 2011). In men, however, only a few studies have reported an own-gender bias (e.g., Wright & Sladden, 2003). These findings were highlighted in Herlitz and Lovén's (2013) meta-analysis which demonstrated that the own-gender bias effect found in women was reliable, but this bias was not displayed in men.

In addition to participant biases, some studies have found that female faces were recognised more accurately than male faces (e.g., Lewin & Herlitz, 2002; Lovén et al., 2011) by both female and male participants (e.g., Rehnman & Herlitz, 2006). Additionally, Rehnman and Herlitz (2007) found that women were generally more accurate at recognising both male and female faces compared to men, demonstrating sex differences favouring women in face recognition. However, the extent with which this sex difference occurs is unclear as several studies have indicated that this finding in women was only present towards female faces, and was not present for male faces when both male and female faces were included in the test material (Lewin & Herlitz, 2002; McKelvie, 1987; Slone, Brigham, & Meissner, 2000). Herlitz and Lovén (2013) proposed that the strong own-gender bias in women might have occurred from an early perceptual interest in female faces (Connellan, Baron-Cohen, Wheelwright, Batki, & Ahluwalia, 2000), which has persisted throughout development through a larger number of intense and important social interactions with other females (such as primary caregivers) relative to males (Rennels & Simmons, 2008). This has led to perceptual expertise for female faces.

Perceptual expertise has been used to account for the own-group biases. Expert processing is typically employed for faces but not employed for objects. This expertise is typically defined as configural encoding (Maurer, Le Grand, & Mondloch, 2002). One method for testing expert processing of faces is with the face-inversion effect (Farah, Tanaka, & Drain, 1995). Studies have continuously demonstrated that the recognition of inverted faces is disproportionately more affected than inverted objects (e.g., Valentine, 1988; Yin, 1969). Inverted faces are typically processed in an inexpert manner (Kanwisher, Tong, & Nakayama, 1998; McKone & Yovel, 2009; Tanaka & Farah, 1993). Reduced expert processing (and a smaller face-inversion effect) has been revealed for faces of other groups relative to own group faces (Hills & Willis, 2016; Hugenberg & Corneille, 2008; Michel, Rossion, Han, Chung, & Caldara, 2006; Rhodes, Brake, Taylor, & Tan, 1989). The perceptual expertise account is typically contrasted with the socio-cognitive motivational accounts. In this, faces are quickly coded by their group status and in-group faces are processed deeply, whereas out-group faces are processed shallowly (Sporer, 2001). One tool that has been used to test between these theories is eye tracking.

Eye-tracking studies have indicated that internal facial features (the eyes, nose, and mouth) are fundamental for face recognition as these features receive more fixations than other features, especially for upright faces (e.g., Althoff & Cohen, 1999; Stacey, Walker, & Underwood, 2005; Williams & Henderson, 2007; Zhao, Chellappa, Phillips, & Rosenfeld, 2003). Henderson, Falk, Minut, Dyer, and Mahadevan (2001) reported that during face recognition tasks, approximately 60% of fixation time was spent on the eyes (Williams & Henderson, 2007), thereby indicating the importance of the eyes when processing faces (Gosselin & Schyns, 2001).

These results have been found with participants viewing faces that one would expect they would be employing expert processing for (own-group upright faces). Previous eye-tracking research in face expertise and familiarity demonstrated that external features are used more often to help process and recognise unfamiliar faces (Stacey et al., 2005). Furthermore, Goldinger, He, and Papesh (2009) found differences in the way that own- and other-ethnicity faces are viewed, with a less effortful encoding strategy highlighted by fewer and longer fixations with more fixations to the nose for other-ethnicity faces relative to own-ethnicity faces (see also, Brielmann, Bülthoff, & Armann, 2014; Fu, Hu, Wang, Quinn, & Lee, 2012; but see, Blais, Jack, Scheepers, Fiset, & Caldara, 2008). Similarly, Hills and Willis (2016) have shown that participants focus more on the nose in other-age faces relative to own-age faces. These results indicate differential scanning for own- and other-group faces. Goldinger et al. (2009) interpret these findings according to the socio-cognitive account of the own-group biases, suggesting that the group status of the face leads to differential processing of the face.

Eye tracking evidence can also reveal expertise in processing faces through greater attention to the most diagnostic features of a face for recognition (Caldara, Zhou, & Miellet, 2010; Kelly, Jack, Miellet, De Luca, Foreman, & Caldara, 2011; Rodger, Kelly, Blais, & Caldara, 2010). Due to extensive experience, participants attend to the most diagnostic features of the faces they most frequently encounter (Hills & Pake, 2013). Perceptual expertise when looking at faces can also be revealed through the number and length of fixations (Bombari, Mast, & Lobmaier, 2009). Expertise in many perceptual domains is associated with a wide perceptual field, such that, with a single central fixation, an entire face can be viewed (Rossion, 2008). Expert processing is therefore revealed through fewer longer fixations to the most diagnostic parts of a face. Indeed, children younger than 10-years of age tend to show a fixation pattern characteristic of inexpert processing: they show more shorter fixations distributed across the whole face (Hills, Willis, & Pake, 2014).

A useful addition to test this expertise account using eye tracking is with the face-inversion effect, since inverted faces are processed using inexpert mechanisms (Tanaka & Farah, 1993). Inversion reduces the size of the perceptual field leading to more (and shorter) fixations required to process a face (Rossion, 2008). Evidence from Barton, Radcliffe, Cherkasova, Edelman, and Intriligator (2006) supports this as they found that participants' scan path was more random when viewing inverted faces relative to upright faces, with more fixations towards less diagnostic features (see also, Hills, Sullivan, & Pake, 2012).

Exploring sex differences in face recognition using eye-tracking, Heisz, Pottruff, and Shore (2013) suggested that potential differences in face recognition due to sex was directly related to their scanning behaviour when processing faces. They found that women made more fixations during initial encoding than men did. However, during initial encoding, there were no sex differences in the distribution of fixations made across the eyes, nose, and mouth. It was only after repeated exposures to the facial stimuli that a sex difference in the distribution of fixations was found whereby women directed significantly more attention towards female eyes compared to male eyes. Similar findings have been shown in expression recognition tasks, with women directing more fixation to the eyes (Hall, Hutton, & Morgan, 2010; Vassallo, Cooper, & Douglas, 2009). Despite replicating findings of a sex difference in face recognition, and giving some insight into the way men and women process faces differently, Heisz et al. (2013) did not specifically report an own-gender bias in their behavioural data occurring for either men or women. As such, there is currently limited research explaining the mechanisms behind the own-gender bias in face recognition. Since Heisz et al. (2013) were not able to assess perceptual expertise in their study, it is hard to establish what the mechanisms of the own-gender bias are in face recognition. Face inversion during eye tracking will allow us to compare whether the own-gender bias is based on perceptual expertise (as indicated by a larger-face inversion effect and more distributed eye movements over faces) or socio-cognitive motivational accounts, in which other-gender faces will be scanned in a less effortful manner than own-gender faces.

The literature summarised here indicates that while the mechanism behind the own-gender bias is largely unclear, it may be based on some form of perceptual expertise or socio-cognitive motivational factors. We devised a study to test the involvement of face expertise in the own-gender bias. To do this, we employed the face-inversion effect as a measure of face expertise. Additionally, we employed eye-tracking in order to test Rossion's (2008) claims regarding expertise. Our primary aim of this study was to investigate sex differences, specifically the own-gender bias, in face recognition using eye-tracking. Based on previous research, we predict an own-gender bias in women but not all men and that women would outperform men in the face recognition task (Heisz et al., 2013; Herlitz & Lovén, 2013). If the own-gender bias is based on perceptual expertise, we would expect to see a larger face-inversion effect and more expert scanning behaviour to own-gender faces than other-gender faces. Behaviourally, we would predict a three-way interaction between participant gender, face gender, and inversion if the perceptual expertise account explained the own-gender bias. If the socio-cognitive motivational account is more applicable, then we would expect to see less effortful scanning of other-gender faces than own-gender faces, thereby a three-way interaction between participant gender, face gender, and feature viewed.

Method

Participants

An opportunity sample of 81 psychology undergraduate students (aged between 18 and 36 years, mean: 21 years, 41 women) from Anglia Ruskin University participated in the study. One female participant was excluded from the experiment as there were difficulties in calibrating the eye-tracker to her eyes. All participants self-reported as being ethnically White from a sample in Cambridge in order to match the face stimuli and thereby avoiding any own-ethnicity bias effects. Additionally, participants self-reported having normal or corrected-to-normal vision. All participants received course credits for their participation.

Apparatus

Ninety-six facial images (48 female and 48 male; aged between 18 and 30 years) were taken from the Nottingham Scans and Aberdeen databases (http://www.pics.stir.ac.uk). All faces were ethnically White and displayed neutral expressions in frontal poses. Stimuli were presented in greyscale on a white background in the centre of a 17″ LCD colour monitor. All facial stimuli were constrained to the proportions 350-pixels by 275-pixels (subtending 11.26 by 8.85 degrees of visual angle) using GNU Image Manipulation Program (GIMP; version 2.8.10). All facial stimuli displayed hairstyles but other distinctive characteristics, such as jewellery, were removed: this was achieved by using the *"resynthesizer"* tool to remove and edit the selected areas. Inverted stimuli were created by rotating the stimuli by 180°. During analysis, five non-overlapping areas of interest (AOIs) were mapped onto each face (as illustrated in Figure 1) in a similar manner as Goldinger et al. (2009): We ran analyses using different AOIs (such as not including the eyebrows within the eye region): All results were broadly consistent with the approach taken here.

Figure 1 about here

A grey box was also created using GIMP, with the same dimensions as the facial stimuli. The purpose of the grey box was to act as a cue (similarly to a fixation cross), while avoiding affecting the validity of the eye-tracking data (Hills & Pake, 2013) by allowing participants to look anywhere within the box region before each face appeared, rather than fixating on a cross at the centre of the screen (see also, Wu, Laeng, & Magnussen, 2012).

E-Prime Professional 2.0 was used to create and run the experiment. ClearView software (version 2.7.0) was used to calibrate the eye-tracker to participants' eyes. An eye-tracker (Tobii 1750; Falls Church, VA, USA) was used to record participants' eye movements. This eye-tracker emits near infra-red light, which is reflected off the observer's eyes. The eye-tracker detects this reflection and computes the location of the fixation relative to the screen. The eye-tracker had a sampling rate of 50 Hz and a spatial resolution of 0.25 degrees of visual angle. All stimuli were presented onto a colour 17" LCD monitor (dimensions 1280 x 1024). A standard chinrest was used to position participants' head and to ensure their heads were kept within the eye-tracker range for the eye-tracker to be calibrated successfully.

Design

A 2 (sex of participants: women, men) x 2 (gender of faces: female, male) x 2 (orientation: upright, inverted) mixed design was used to measure recognition accuracy, using Signal Detection Theory (*d'*; Green & Swets, 1966). The eye-tracking measure total duration of fixation to each AOI was recorded. Sex of participants was a between-subjects variable, whilst gender of faces and orientation were manipulated within-subjects. Faces were presented in a random order with an equal number of female and male faces presented upright and inverted during both phases of the experiment. Faces were counterbalanced such that they each appeared as a distractor and a target an equal number of times.

Procedure

Participants were tested individually in a quiet laboratory. Once participants gave their informed consent, they were seated approximately 60 cm from the computer screen and had their head comfortably placed on a standard chinrest to minimise head movements. Using ClearView software, participants' eyes were calibrated to the eye-tracker: participants were required to follow a moving ball to five pseudo-random locations on-screen while the eye-tracker recorded their eye movements. The eye-tracker was recalibrated until an accurate and reliable measure of the estimated eye position was achieved.

Once calibrated, participants were asked to answer two demographic questions on the computer, regarding their age and gender. Following this, a standard old/new face recognition paradigm was employed involving a learning, break, and test phase.

For the learning task, 48 faces (24 females and 24 males, half of each inverted) were presented sequentially in the centre of the screen. A grey box was presented in the centre of the screen for 500 ms before the face appeared in the same place as the face would appear. Faces were on screen for 2,000 ms during which time participants were required to rate their distinctiveness using a 7-point scale, using the numerical keypad (1 indicated "Not at all" and 7 indicated "Very"). Participants were verbally informed that they could use the numbers in between as well. Rating the facial stimuli ensured that the participants attended to each face.

The eye-tracker was recalibrated to ensure that an accurate estimation of participants' eye movements was recorded for the recognition task. This served as a break and typically lasted 45 seconds. For the recognition task, participants stated whether they had seen each face before in the learning phase. They were informed that the accuracy of their performance were more important than their response speed. Participants were presented with 96 faces sequentially in a random order: 48 were old ‘target’ faces that had appeared during the learning task while the remaining 48 were new ‘distractor’ faces (half upright and half inverted). The target faces were presented in the same orientation as they were presented in the learning task. The same number of female and male faces in each orientation for the target faces was used for the new faces as old faces. A grey box presented for 500 ms preceded each face. Each face remained on screen until the participant responded. Once all the stimuli were presented, participants were thanked and verbally debriefed. The duration of the experiment was approximately 20 minutes.

Results

For clarity, we shall present the behavioural data separately from the eye-tracking data. We first wanted to establish that any effects observed in accuracy and in eye-movements were not due to inherent differences in the distinctiveness of the stimuli, followed by an exploration of the recognition performance of our participants.

Learning behavioural data

A 2 (sex of participants) x 2 (gender of faces) x 2 (orientation) mixed factorial ANOVA was used to analyse face distinctiveness. Results (summarised in Table 1) indicate a significant effect of orientation, whereby upright faces (*M*=3.33, *SE*=0.12) were reported as being more distinctive than inverted faces (*M*=3.09, *SE*=0.11), $F(1,78)=7.16$, $MSE=0.64$, $p<.01$, $\eta_g^2=.01$. Additionally, there was a significant interaction between gender of faces and sex, $F(1,78)=5.82$, $MSE=0.68$, $p<.05$, $\eta_g^2=.01$. Bonferroni-corrected post-hoc paired *t*-tests were conducted to explore this interaction: There was a trend for women (*M=*3.46, *SE=*0.18) to rate female faces as being more distinctive than men did (*M=*2.99, *SE=*0.15), $t(39)=1.87$, $p=.07$, but they did not rate male faces (*M=*3.17, *SE*=0.15) as more distinctive than men did (*M*=3.01, *SE*=0.15), $t(39)=1.62$, $p=.114$. Critically, the main effect of face gender was not significant, $F(1, 78)=0.14$, $MSE=0.68$, $p=.71$. Therefore, we can accept any following effects are not due to differences in distinctiveness in the stimuli. There were no further significant main effects nor interactions, largest *F*(1,78)=0.77, smallest *p*=.38.

Recognition behavioural data

A parallel 2 x 2 x 2 mixed ANOVA was used to analyse potential sex differences in mean recognition accuracy (*d'*) for male and female faces in upright and inverted positions (as illustrated in Table 1). Mean recognition accuracy was calculated using Macmillan and Creelman's (2005) method of Signal Detection Theory (*d'*; Green & Swets, 1966): Responses from the old and new faces were converted into hit rates and false alarm rates and then into *d'*. All 0 counts were replaced with 0.5 (Macmillan & Creelman, 2005). Two participants' scores revealed outliers (scores above 1.96 *SD*s away from the mean) in two conditions: these values were replaced by means for that condition.

Overall recognition performance for men and women did not differ significantly, $F(1,78)<0.01$, $MSE=0.73$, $p=.95$, $\eta_g^2<.01$. This is inconsistent with Herlitz and Lovén (2013). However, there was a significant main effect for gender of faces, $F(1,78)=11.99$, $MSE=0.30$, $p=.001$, $\eta_g^2=.03$, whereby female faces ($M=1.22$, $SE=0.06$) were recognised more accurately than male faces ($M=1.01$, $SE=0.05$). These findings were observed due to the predicted (therefore a one-tailed test) interaction between face gender and participant sex being significant, $F(1,78)=2.74$, $MSE=0.33$, $p=.05$, $\eta_g^2=.01$. Bonferroni-corrected *t*-tests revealed that for women, the recognition of female faces ($M=1.27$, $SE=0.10$) was greater than the recognition of male faces ($M=0.96$, $SE=0.08$), $t(39)=3.20$, $p=.003$ for upright faces. Men did not show the own-gender bias, $t(39)=1.50$, $p=.141$: the recognition of female faces ($M=1.16$, $SE=0.08$) was statistically not different to the recognition of male faces ($M=1.05$, $SE=0.07$). This is consistent with Lovén et al. (2011).

In order to assess whether own-gender faces were processed in an expert manner than other-gender faces, we looked at the three-way interaction between face gender, participant sex, and orientation of the face. This was not significant, $F(1,78)=0.20$, $MSE=0.32$, $p=.655$, $\eta_g^2<.01$. Additionally, women did not show a significant own-gender bias for inverted faces, $t(39) = 1.74$, $p = .09$, nor did men, $t(39) = 0.63$, $p = .53$. The main effect of orientation was significant overall, however, $F(1,78)=122.54$, $MSE=0.42$, $p<.001$, $\eta_g^2=.27$, whereby upright faces ($M=1.51$, $SE=0.06$) were recognised more accurately than inverted faces ($M=0.71$, $SE=0.06$).

Table 1 about here

Eye-tracking data

We present the eye-tracking data during the learning phase as the effect we are theoretically interested in is that of more depth of encoding given the results of the behavioural data. Data analysis from the recognition phase revealed a similar pattern to those found in the recognition analysis (including phase of the experiment into the analysis did not reveal any significant interactions with phase, all *p*s>.35), and are available from the corresponding author. During the analysis, any trial in which a face was fixated on for less than 200 ms was removed (this occurred in less than 1% of trials for any participant and occurred randomly across conditions). Using the standard defaults of the Tobii eye-tracker, a fixation was defined as an eye being recorded as looking at the same 30 pixel region for at least 100 ms (see Goldinger et al., 2009). Any eye movement that left an AOI but returned to it within 100 ms was considered the same fixation. Mean total duration of fixation (ms) as defined as the total amount of time a particular AOI was fixated upon (see Goldinger et al., 2009). Mauchly's test of sphericity indicated that the assumptions of sphericity were violated for all analyses involving AOI; therefore, the *F*-values were corrected using Greenhouse-Geisser correction. Here, we present the results as directed by each of our hypotheses.

Duration of fixation

We first ran a 2 x 2 x 5 (sex of participants x gender of faces x AOI) mixed ANOVA on the total time fixating on each AOI (the data produced a similar pattern to the total number of fixations to each AOI data) exploring only the upright faces. Means are presented in Figure 2a. Firstly, we found that female faces (*M*=388, *SE*=1.46) were looked at significantly longer than male faces (*M*=381, *SE*=1.49), $F(1, 78)=15.11$, *MSE*=598.28, $p<.001$, $\eta_g^2=.02$, though this effect interacted with participant sex, $F(1, 78)=4.16$, *MSE*=598.28, $p=.045$, $\eta_g^2=.01$. Women fixated more at female faces (*M*=1951ms, *SE*=9.90) than male faces (*M*=1900ms, *SE*=11.59), $t(39)=4.28$, $p<.001$. Men did not fixate significantly more at female faces (*M*=1929ms, *SE*=10.77) than male faces (*M*=1913ms, *SE*=9.29), $t(39)=1.28$, $p=.208$.

The gender of the face interacted with the AOI, $F(4, 312)=16.72$, *MSE*=1385277, $p<.001$, $\eta_g^2=.02$. For female faces (*M*=448ms, *SE*=30.42), the forehead/hair AOI was fixated on more than for male faces (*M*=326ms, *SE*=27.98), $t(79)=6.24$, $p<.001$, indicating the importance of the hair when processing female faces (see e.g., Wright & Sladden, 2003). Moreover, the nose was fixated on more in male faces (*M*=427ms, *SE*=22.43) than for female faces (*M*=306ms, *SE*=20.66), $t(79)=3.70$, $p<.001$.

We found the standard hierarchy of features, revealed through the main effect of AOI, $F(4, 312)=186.26$, *MSE*=159932.00, $p<.001$, $\eta_g^2=.64$. Šidák-corrected post-hoc paired comparisons indicated that the eyes received more fixation that every other features, the nose and forehead/hair received more fixation than the mouth and chin areas (smallest mean difference = 234 ms, *SE*=22.15, largest $p<.001$), though there was no fixation difference across the nose and forehead/hair (mean difference=21 ms, p =1.00) nor the mouth and chin areas (mean difference=18ms, $p=.996$).

Figure 2 about here

We next looked at the interaction between participant sex and AOI to examine whether there were any differences in the way men and women processed faces in general. This was not significant, $F(4, 312)=1.26$, $MSE=159932$, $p=.286$, $\eta_g^2=.01$. Specifically, women (M=907ms, SE=35.75) did not look significantly more at the eyes than men (M=937ms, SE=44.90), $t(78)=0.52$, $p=.604$, and men (M=347ms, SE=31.03) did not look significantly more at the nose than women (M=385ms, SE=17.27), $t(78)=1.08$, $p=.286$. These results are inconsistent with Hall et al. (2010) and Vassallo et al.'s (2009) findings.

To directly test the socio-cognitive account of the own-gender bias, we looked at the three-way interaction between participant sex, face gender, and AOI. This was significant, $F(4, 312)=48.66$, $MSE=28468$, $p<.001$, $\eta_g^2=.09$. Šidák-corrected ($\alpha=.005$) post-hoc paired *t*-tests between male and female faces in each AOI for both men and women revealed that: both men and women looked at the eyes more in own-gender faces, $t(39)=6.54$, $p<.001$ (men) and $t(39)=6.26$, $p<.001$ (women). Men and women looked at the hair in female faces more than in male faces, $t(39)=3.486$, $p=.001$ (men) and $t(39)=6.28$, $p<.001$ (women). Finally, women looked at the nose more in male faces than female faces, $t(39)=9.25$, $p<.001$; this was a non-significant trend after Šidák -correction for men, $t(39)=2.08$, $p=.044$.

The Effect of Inversion

We then ran a parallel analysis on the inverted faces. Given the results from the behavioural analysis, we were not expecting a gender bias in the eye movements over inverted faces as there was no gender bias observed in the accuracy with which inverted faces were recognised (see above analysis). This analysis did not reveal any interactions between participant sex and any variables, largest $F=1.16$, smallest $p=.323$, largest $\eta_g^2<.01$, as indicated by the means shown in Figure 2b. There was, however, a main effect of AOI, $F(4, 312)=62.85$, $MSE=329441$, *p*<.001, η_g^2=.01. All features were viewed a significantly different amount to each other (smallest mean difference=45ms, *p*=031). The hierarchy of features was such that the chin was scanned more than the eyes, followed by the forehead, nose, then finally the mouth.

By combining the two analyses into a 2 x 2 x 2 x 5 (participant sex by face gender by orientation by AOI) mixed ANOVA, we were also able to assess Rossion's (2008) theory that inversion affects scanning behaviour by reducing the perceptual field by assessing the interaction between AOI and facial orientation. This was significant, $F(4, 312)=119.48$, *MSE*=184133, *p*<.001, η_g^2=.10. Šidák-corrected post-hoc paired *t*-tests revealed that significantly more time was spent fixating on the eyes, and nose in upright faces compared to inverted faces, *t*(79)=13.53, *p*<.001 (eyes) and *t*(79)=9.18, *p*<.001 (nose). Significantly more time was spent fixating on the chin in inverted faces relative to upright faces, *t*(79)=15.90, *p*<.001. There were no significant differences in the fixation pattern across upright and inverted faces to the mouth, *t*(79)=1.60, *p*<.001, nor the hair, *t*(79)=0.96, *p*<.001. Orientation did not interact with any of the effects described in the analysis for upright faces.

Discussion

This study used eye-tracking to examine sex differences in face recognition, specifically the own-gender bias in face recognition. We found an asymmetrical own-gender bias effect: women recognised female faces more accurately than male faces (e.g., Lovén et al., 2011; Rehnman & Herlitz, 2007), whereas males did not show an own-gender bias. Analysis from the behavioural data indicated that distinctiveness did not contribute to accuracy performance: while there was a marginal own-gender bias in distinctiveness ratings, distinctiveness did not correlate with hit rate for female participants, *r*(38)=.14, *p*>.05. Our eye-tracking data supported the results regarding the own-gender bias. We found sex differences in scanning movements when processing faces of each group. We found that women looked at the eyes of own-gender faces longer than other-gender faces, similar to Heisz et al.'s (2013) findings. The same pattern was also found for men (possibly due to the fact that 30% of our male participants did show an own-gender bias). In addition to this, women fixated on male noses more than female noses. These results follow similar principles as Goldinger et al.'s (2009) work on the own-ethnicity bias and Hills and Willis' (2016) work on the own-age bias, showing that participants make more fixations to the nose in other-group faces.

Contrary to Heisz et al.'s (2013), we did not find any difference in recognition performance between men and women. The reason for there not being a sex difference could be supported by our eye-tracking data as there was no significant interaction between participant sex and AOI. This contrasts with Heisz et al.'s (2013) study as they found that women outperformed men and their eye-tracking data provided support to their behavioural data explaining that men made fewer fixations compared to women when faces were initially encoded; they suggested that less perceptual information was obtained when they initially processed faces, which meant that they were unable to establish a sufficient memory representation of faces, resulting in poorer performance. Our eye-tracking data supports our behavioural data in that no significant interaction between participant sex and AOI was found i.e., there were no significant difference in the way men and women processed faces (in terms of the importance of features they looked at) to contribute to a better recognition performance.

In our study, we measured the face-inversion effect. Assuming the own-gender bias was based on perceptual expertise in a similar manner to the own-ethnicity bias (e.g., Michel et al., 2006), we would have expected to find an interaction between face gender, participant gender, and orientation (as the face-inversion effect should be larger for own-gender faces than other-gender faces). This was not found in our study. The eye-tracking data support this claim. Men and women were scanning inverted faces in a similar manner and were equally affected by the inversion. While face inversion disrupted performance (replicating the well-established face-inversion effect, Yin, 1969) and eye movements (with shorter more dispersed fixations to inverted faces, replicating Barton et al., 2006 and Hills et al., 2012 work, consistent with Rossion's (2008) theory that inversion restricts the perceptual field), it affected all faces equally. This indicates that the mechanism for the own-gender bias is not based on perceptual expertise and increased use of holistic processing. Furthermore, as mentioned in our introduction, there has not been any previous research incorporating face inversion into an eye-tracking study to examine sex differences in face recognition so this hypothesis was based on research studying cultural differences, where an own-ethnicity bias effect was found. This result contrasts the own-gender bias with the own-age and the own-ethnicity biases.

In the introduction, we presented two broad models for the own-group biases in face recognition. One was the perceptual expertise account, in which, we would expect to see sex differences when scanning faces and an interaction between facial orientation, face gender, and participant sex. Neither were observed, indicating that the mechanism for the own-gender bias is not based on perceptual expertise and increased use of holistic processing. Instead, it might be based on the second model we presented in the introduction: the socio-cognitive model (e.g., Sporer, 2001). In this, participants are presented with a face and immediately make a category judgement. Own-group faces are then processed deeply and other-group faces are processed shallowly. This should be revealed by differential scanning when looking at own- and other-group faces (cf., Goldinger et al., 2009). This is what we have found here: our female participants, who displayed an own-gender bias, looked at the eyes of their own-gender faces more than other-gender faces (consistent with findings on the own-ethnicity bias, Goldinger et al., 2009). By not coding other-gender faces appropriately (not viewing the appropriate features and making fewer fixations to the facial features) leads to poorer recognition performance. In addition to this, both men and women looked at the hair/forehead more in female faces than males, suggesting that they are both using external perceptual information to help them make a categorical judgement. As the eyes have been known to be the most fundamental information needed for face recognition (Gosselin & Schyns, 2001), both men and women were able to identify and extract vital perceptual information to form a memory representation of the faces. However, this account may not be sufficient enough to explain our results because our findings indicate that only women displayed an own-gender bias: according to this theory, men should have processed male faces at a deeper level as they have been categorised as an own-group face, thus, men should have also displayed an own-gender bias.

To resolve this discrepancy, we can return to the account of Herlitz and Lovén (2013) who suggest that female faces are exposed to men and women at a greater extent during infancy than male faces as quite often their primary (i.e., mother) and secondary caregivers (i.e., teachers) are female. This leads female faces to be preferentially processed by both males and females (Rehman & Herlitz, 2006). This preferential processing potentially led to deeper processing and led to female faces being recognised more accurately than males faces in the current study, consistent with Lewin and Herlitz (2002; Lovén et al., 2011). Studies have demonstrated that children engage and interact more with their own-sex peers than with other-sex peers (Maccoby, 1990), which appears to persist as children grow older (Martin, Fabes, Evans & Wyman, 1999). This may therefore provide greater opportunities for girls to enhance their perceptual advantage towards female faces, which can further explain why women displayed an own-gender bias. Linking this to the socio-cognitive account, women's expertise towards female faces might be further enhanced as they might categorise them as an in-group member and thus would process them at a much deeper level. Men, however, may not have displayed an own-gender bias for the very reason that despite categorising male faces as an in-group they are unable to process them at a deeper level due to perhaps a balanced or greater exposure to female faces at such a young age.

The explanation for the present data is that participants have greater motivation to process own-gender faces (and there is more exposure to female faces developmentally). This leads them to be processed more efficiently (Sporer, 2001). This greater motivation leads to the internal features (and in particular the eyes) to be processed more efficiently. The processing of internal facial features accurately is important for two reasons. As demonstrated in Stacey et al.'s (2005) face-matching study, internal features play an important role during the processing of faces, regardless of the familiarity of faces. Furthermore, Stacey et al. (2005) found greater reliance on the external features when processing unfamiliar faces, suggesting that external features provided additional information that help decide whether faces were previously seen or not. As such, our study may build upon the framework of previous research that have examined eye movement patterns when processing unfamiliar faces, enabling us to gain some understanding of the saliency of internal and external features when men and women process unfamiliar faces. Specifically, we have shown the importance of hair when processing female faces (see also Wright & Sladden, 2003), as both men and women fixated more on female hair and used it as a tool for face identification, potentially highlighting the utility of this feature for distinguishing unfamiliar female faces more readily than male faces.

A second reason for the importance of processing the internal features (particularly the eyes) is that eye-tracking evidence has indicated that fixations to the centre of the face lead to greater holistic processing of faces (Blais et al., 2008). However, our behavioural results regarding the effects of inversion tend to indicate that own-gender faces were not processed more holistically than other-gender faces (assuming inversion is a test of holistic processing). This may indicate that the internal feature processing might actually reflect greater depth of processing rather than an increase in holistic processing. As both men and women fixated on the eyes of their own-gender faces more than the other-gender faces, our results might suggest that both sexes engage in holistic processing for own-gender faces more so than other-gender faces.

The reason why female participants then show the behavioural own-gender bias more so than male participants might be that additional perceptual expertise or motivation gained through extensive experience throughout development is also needed to strengthen the ability to combine these features into a single representation.

In summary, we found an own-gender bias in female participants but not male participants. Our eye-tracking analysis indicated that own- and other-gender faces are scanned in a different way, with more fixations to the eyes in own-gender faces. Given the lack of a larger face-inversion effect in own-gender faces relative to other-gender faces, and these scanning differences, we explain these results within a socio-cognitive account of the own-gender bias, whilst acknowledging greater experience with female faces developmentally (Rennels & Simons, 2008) leading to women demonstrating an own-gender bias.

References

Althoff, R., & Cohen, N. (1999). Eye-movement-based memory effect: A reprocessing effect in face perception. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(4), 997–1010. doi: 10.1037/0278-7393.25.4.997

Barton, J. J. S., Radcliffe, N., Cherkasova, M. V., Edelman, J., & Intriligator, J. M. (2006). Information processing during face recognition: The effects of familiarity, inversion, and morphing on scanning fixations. *Perception, 35*, 1089-1105. doi:10.1068/p5547

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS One, 3*, e3022. doi:10.1371/journal.pone.0003022

Bombari, D., Mast, F. W., & Lobmaier, J. S. (2009). Featural, configural, and holistic face-processing strategies evoke different scan patterns. *Perception*, *38,* 1508-1521.

Brielmann, A. A., Bülthoff, I., & Armann, R. (2014). Looking at faces from different angles: Europeans fixate different features in Asian and Caucasian faces. *Vision Research*, *100*, 105-112. doi:10.1016/j.visres.2014.04.011

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77,* 305-327. doi:10.1111/j.2044-8295.1986.tb02199.x

Caldara, R., Zhou, X., & Miellet, S. (2010). Putting culture under the ‘spotlight’ reveals universal information use for face recognition. *PLoS One*, *5*, e9708. doi:10.1371/journal.pone.0009708

Chiroro, P., & Valentine, T. (1995). An investigation of the contact hypothesis of the own-race bias in face recognition. *The Quarterly Journal of Experimental Psychology*, *48,* 879-894. doi:10.1080/14640749508401421

Connellan, J., Baron-Cohen, S., Wheelwright, S., Batki, A., & Ahluwalia, J. (2000). Sex differences in human neonatal social perception. *Infant Behavior and Development*, *23*, 113-118. doi:10.1016/S0163-6383(00)00032-1

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect?. *Journal of Experimental Psychology: Human perception and performance*, *21*(3), 628.

Feingold, G. A. (1914). Influence of environment on identification of persons and things. *Journal of Criminal Law and Criminology, 5*, 39-51

Fu, G., Hu, C. S., Wang, Q., Quinn, P. C., & Lee, K. (2012). Adults scan own-and other-race faces differently. *PLoS One, 7*(6), e37688. doi:10.1371/journal.pone.0037688

Goldinger, S. D., He, Y., & Papesh, M. H. (2009). Deficits in cross-race face learning: Insights from eye movements and pupillometry. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(5), 1105-1122. doi:10.1037/a0016548

Gosselin, F., & Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision Research, 41*, 2261-2271

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Hall, J. K., Hutton, S. B., & Morgan, M. J. (2010). Sex differences in scanning faces: Does attention to the eyes explain female superiority in facial expression recognition?. *Cognition and Emotion*, *24*, 629-637. doi:10.1080/02699930902906882

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in cognitive sciences, 4*, 223-233. doi:10.1016/S1364-6613(00)01482-0

Heisz, J. J., Pottruff, M. M., & Shore, D. I. (2013). Females scan more than males: A potential mechanism for sex differences in recognition memory. *Psychological Science*, *24*, 1157-1163. doi:10.1177/0956797612468281

Herlitz, A., & Lovén, J. (2013). Sex differences and the own-gender bias in face recognition: A meta-analytic review. *Visual Cognition*, *21*, 1306-1336. doi:10.1080/13506285.2013.823140

Hills, P. J., & Pake, J. M. (2013). Eye-tracking the own-race bias in face recognition: Revealing the perceptual and socio-cognitive mechanisms. *Cognition*, *129*, 586-597. doi:10.1016/j.cognition.2013.08.012

Hills, P. J., Sullivan, A. J., & Pake, J. M. (2012). Aberrant first fixations when looking at inverted faces in various poses: The result of the centre-of-gravity effect?. *British Journal of Psychology*, *103*, 520-538. doi:10.1111/j.2044-8295.2011.02091.x

Hills, P. J., & Willis, S. F. (2016). Children view own-age faces qualitatively differently to other-age faces. *Journal of Cognitive Psychology*, *28*(5), 601-610.

Hills, P. J., Willis, S. F., & Pake, J. M. (2013). Eye-tracking face recognition in children: effects of age, familiarity, and orientation. Presented at the BPS Joint Cognitive and Developmental Sections Conference, Reading, 4th-6th September 2013.

Hugenberg, K., & Corneille, O. (2009). Holistic Processing Is Tuned for In-Group Faces. *Cognitive Science*, *33*(6), 1173-1181.

Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition*, *68*(1), B1-B11.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences, 361*, 2109-2128. doi:10.1098/rstb.2006.1934

Kelly, D. J., Jack, R. E., Miellet, S., De Luca, E., Foreman, K., & Caldara, R. (2011). Social experience does not abolish cultural diversity in eye movements. Frontiers in Psychology, 2. doi:10.3389/fpsyg.2011.00095

Lewin, C., & Herlitz, A. (2002). Sex differences in face recognition – Women's faces make the difference. *Brain and Cognition*, *50*, 121-128. doi:10.1016/S0278-2626(02)00016-7

Lovén, J., Herlitz, A., & Rehnman, J. (2011). Women's own-gender bias in face recognition memory: The role of attention at encoding. *Experimental Psychology*, *58*, 333-340. doi:10.1027/1618-3169/a000100

Maccoby, E. E. (1990). Gender and relationships: A developmental account. *American Psychologist*, *45*, 513-520.

MacMillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide*. New York: Cambridge University Press

Martin, C. L., Fabes, R. A., Evans, S. M., & Wyman, H. (1999). Social cognition on the playground: Children's beliefs about playing with girls versus boys and their relations to sex segregated play. *Journal of Social and Personal Relationships, 16*, 751-771. doi: 10.1177/0265407599166005

McKelvie, S. J. (1987). Sex differences, lateral reversal, and pose as factors in recognition memory for photographs of faces. *The Journal of General Psychology*, *114*, 13-37. doi:10.1080/00221309.1987.9711052

McKone, E., & Yovel, G. (2009). Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? Toward a new theory of holistic processing. *Psychonomic Bulletin & Review, 16*(5), 778-797.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7,* 3-35. doi:10.1037//1076-8971.7.1.3

Michel, C., Rossion, B., Han, J., Chung, C. S., & Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychological Science, 17*(7), 608-615.

O'Toole, A. J., Deffenbacher, K. A., Valentin, D., & Abdi, H. (1994). Structural aspects of face recognition and the other-race effect. *Memory & Cognition*, *22*(2), 208-224

Psychological Image Collection at Stirling (PICS). (2000). Retrieved from http://www.pics.stir.ac.uk

Rehnman, J., & Herlitz, A. (2006). Higher face recognition ability in girls: Magnified by own-sex and own-ethnicity bias. *Memory*, *14*, 289-296. doi:10.1080/09658210500233581

Rehnman, J., & Herlitz, A. (2007). Women remember more faces than men do. *Acta Psychologica*, *124*, 344-355. doi:10.1016/j.actpsy.2006.04.004

Rennels, J. L., & Simmons, R. E. (2008). Facial experience during the first year. *Infant Behavior and Development*, *31*, 665-678. doi:10.1016/j.infbeh.2008.04.009

Rhodes, M. G., & Anastasi, J. S. (2012). The own-age bias in face recognition: A meta-analytic and theoretical review. *Psychological Bulletin*, *138*, 146-174. doi:10.1037/a0025750

Rodger, H., Kelly, D. J., Blais, C., & Caldara, R. (2010). Inverting faces does not abolish cultural diversity in eye movements. *Perception*, *39*, 1491-1503. doi:10.1068/p6750

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica*, *128*, 274-289. doi:10.1016/j.actpsy.2008.02.003

Slone, A. E., Brigham, J. C., & Meissner, C. A. (2000). Social and cognitive factors affecting the own-race bias in whites. *Basic and Applied Social Psychology*, *22*, 71-84. doi:10.1207/S15324834BASP2202_1

Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law*, *7*(1), 36-97. doi:10.1037/1076-8971.7.1.36

Stacey, P. C., Walker, S., & Underwood, J. D. M. (2005). Face processing and familiarity: Evidence from eye-movement data. *British Journal of Psychology*, *96*, 407-422. doi:10.1348/000712605X47422

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology*, *46*, 225-245. doi:10.1080/14640749308401045

Vassallo, S., Cooper, S. L., & Douglas, J. M. (2009). Visual scanning in the recognition of facial affect: Is there an observer sex difference?. *Journal of Vision*, *9*, 1-10. doi:10.1167/9.3.11.

Williams, C. C., & Henderson, J. M. (2007). The face inversion effect is not a consequence of aberrant eye movements. *Memory & Cognition*, *35*, 1977-1985. doi:10.3758/BF03192930

Wright, D. B., Boyd, C. E., & Tredoux, C. G. (2003). Inter-racial contact and the own-race bias for face recognition in South Africa and England. *Applied Cognitive Psychology*, *17*, 365-373. doi:10.1002/acp.898

Wright, D. B., & Sladden, B. (2003). An own gender bias and the importance of hair in face recognition. *Acta Psychologica*, *114*, 101-114. doi:10.1016/S0001-6918(03)00052-0

Wu, E. X. W., Laeng, B., & Magnussen, S. (2012). Through the eyes of the own-race bias: Eye-tracking and pupillometry during face recognition. *Social Neuroscience, 7,* 202-216. doi:10.1080/17470919.2011.596946

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, *79*, 471-491. doi:10.1111/j.2044-8295.1988.tb02747.x

Valentine, T., & Endo, M. (1992). Towards an exemplar model of face processing: The effects of race and distinctiveness. *The Quarterly Journal of Experimental Psychology*, *44*(4), 671-703.

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

Zhao, W., Chellappa, R., Phillips, P. J., & Rosenfeld, A. (2003). Face recognition: A literature survey. *ACM Computing Surveys*, *35*(4), 399-458

Table 1. Mean distinctiveness and recognition accuracy for male and female upright and inverted faces, split by sex of participant. (Standard errors of the mean are presented in parentheses).

		Face orientation and gender			
		Upright		Inverted	
		Male	Female	Male	Female
Distinctiveness Ratings	Men	3.26 (0.17)	3.07 (0.19)	3.10 (0.17)	2.92 (0.17)
	Women	3.35 (0.17)	3.62 (0.19)	3.05 (0.17)	3.29 (0.17)
Recognition Accuracy	Men	1.44 (0.10)	1.58 (0.11)	0.69 (0.10)	0.75 (0.11)
	Women	1.32 (0.10)	1.71 (0.11)	0.60 (0.10)	0.83 (0.11)

Figure 1. Examples of the AOIs used for upright and inverted faces: eyes; nose; mouth; hair (including forehead); and chin.

Figure 2. Mean total fixation duration to each AOI in milliseconds, split by gender of the face and participant sex: (a) Top panel shows the results for upright faces and (b) bottom panel shows the results for the inverted faces. Error bars represent standard error of the mean. * denotes significant different in fixation during to that AOI for male and female faces.

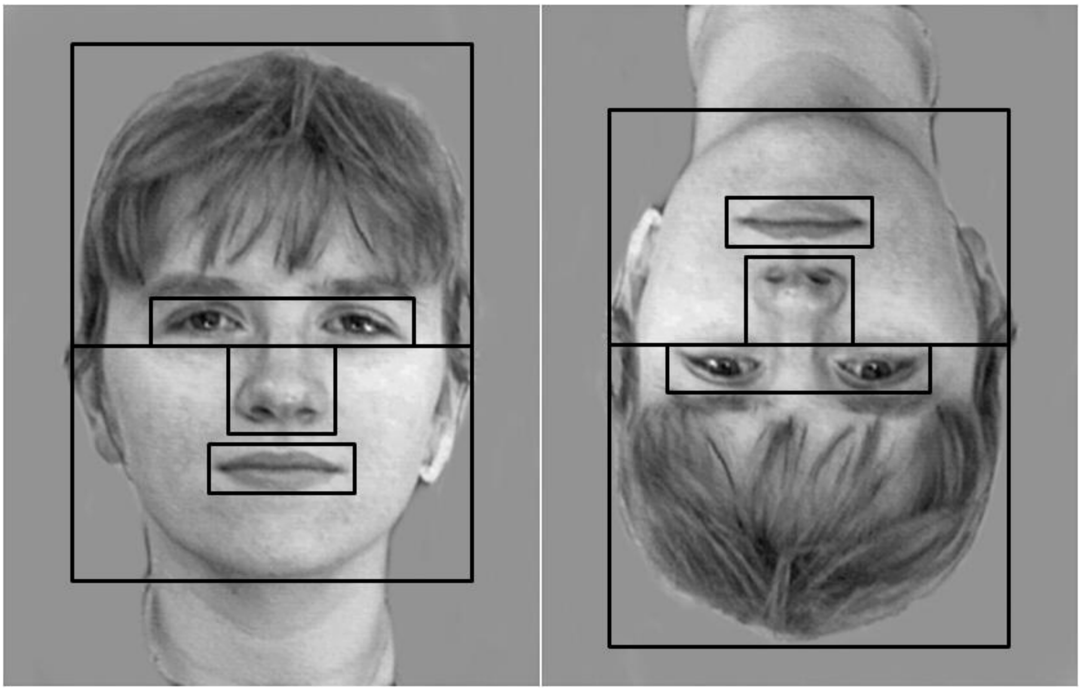
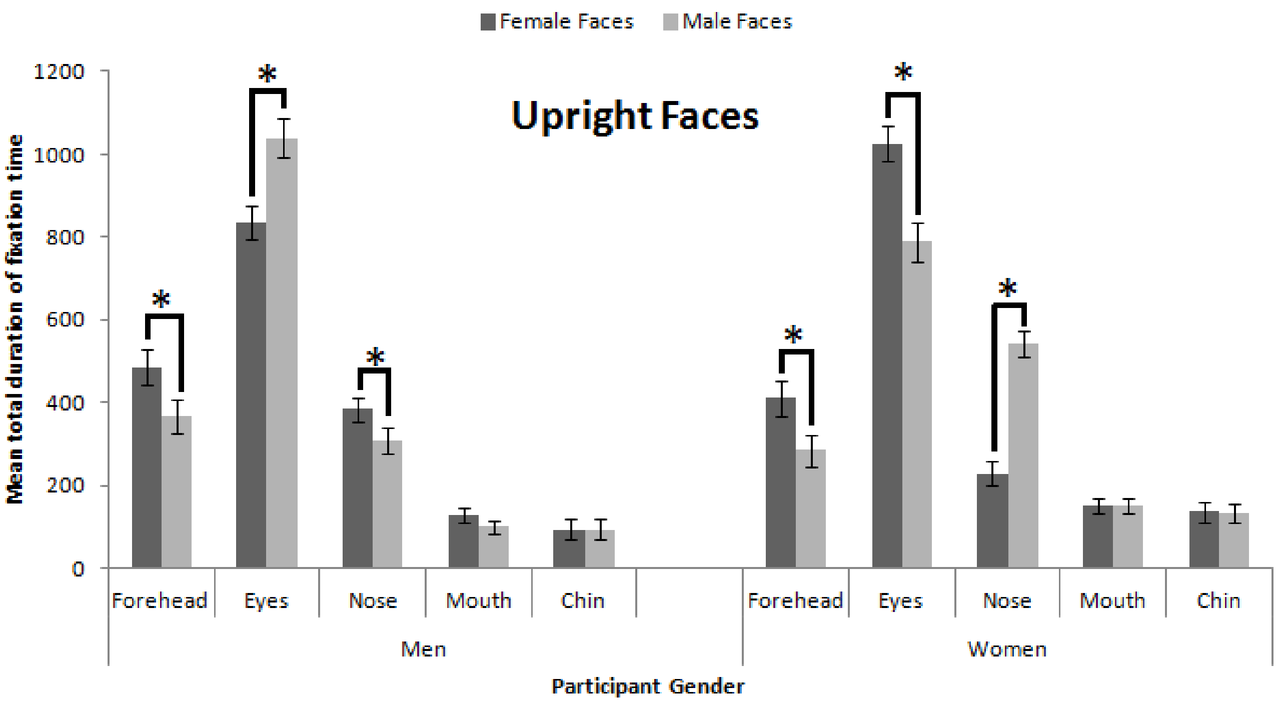
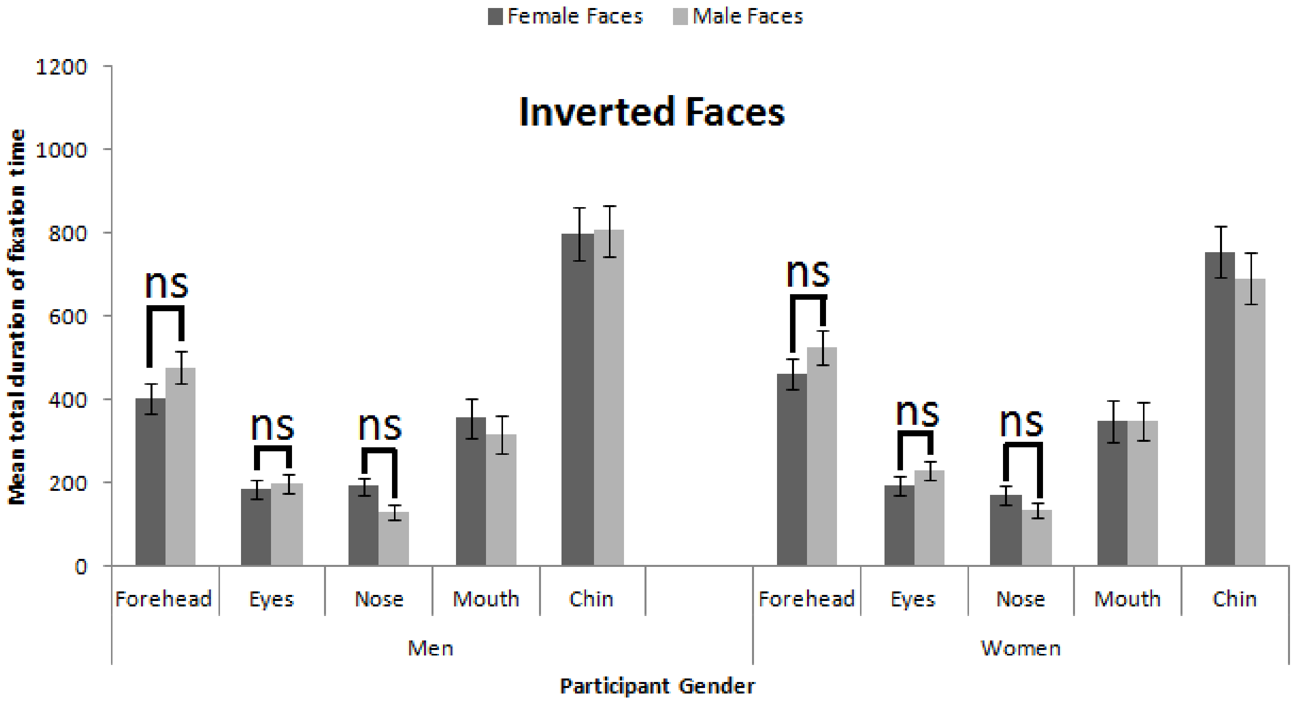


Figure 1.



Female Faces
Male Faces
Upright Faces
Mean total duration of fixation time
0
200
400
600
800
1000
1200
*
Forehead
Eyes
Nose
Mouth
Chin
Men
Women
Participant Gender

a.



Female Faces
Male Faces
Inverted Faces
Mean total duration of fixation time
0
200
400
600
800
1000
1200
ns
Forehead
Eyes
Nose
Mouth
Chin
Men
Women
Participant Gender

b.

Figure 2.