Face exploration dynamics differentiate men and women

Antoine Coutrot
CoMPLEX, University College London, United Kingdom

Nicola Binetti
Dept of Experimental Psychology, University College London, United Kingdom

Charlotte Harrison
Dept of Experimental Psychology, University College London, United Kingdom

Isabelle Mareschal
Queen Mary University of London, United Kingdom

Alan Johnston
School of Psychology, University of Nottingham, United Kingdom

The human face is central to our everyday social interactions. Recent studies have shown that while gazing at faces, each one of us has a particular eye-scanning pattern, highly stable across time. Although variables such as culture or personality have been shown to modulate gaze behaviour, we still don't know what shapes these idiosyncrasies. Moreover most previous observations rely on static analyses of small-sized eye-position datasets averaged across time. Here, we probe the temporal dynamics of gaze to explore what information can be extracted about the observers and what is being observed. Controlling for any stimuli effect, we demonstrate that amongst many individual characteristics, the gender of both the participant (gazer) and the person being observed (actor) are the factors that most influence gaze patterns during face exploration. We record and exploit the largest set of eye tracking data (405 participants, 58 nationalities) from participants watching videos of another person. Using novel data-mining techniques, we show that female gazers follow a much more exploratory scanning strategy than males. Moreover, female gazers watching female actresses look more at the eye on the left side. These results have strong implications in every field using gaze-based models, from computer-vision to clinical

doi: Received: July 26, 2016 ISSN 1534–7362 © 20?? ARVO

psychology.

Keywords: visual exploration, eye-tracking, face perception, Markov models, scanpaths, gender difference

Introduction

Our eyes constantly move around to place our high-resolution fovea on the most relevant visual information. Arguably, one of the most important objects of regard is another person's face. Until recently, a majority of face perception studies have been pointing to a "universal" face exploration pattern: humans systematically follow a triangular scanpath (sequence of fixations) over the eyes and the mouth of the presented face (Yarbus, 1965; Vatikiotis-Bateson, Eigsti, Yano, & Munhall, 1998). However, more recent studies found that face-scanning strategies depend upon many factors, including task (Borji & Itti, 2014; Borji, Lennartz, & Pomplun, 2015), social context (Foulsham, Cheng, Tracy, Henrich, & Kingstone, 2010; Gobel, Kim, & Richardson, 2015), emotion (Eisenbarth & Alpers, 2011; Schurgin et al., 2014), personality (Perlman et al., 2009; Peterson & Eckstein, 2013) and even culture (Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Wheeler et al., 2011). For instance, when watching videos of faces, observers believing that the targets would later be looking at them looked proportionally less at the eyes of the targets with higher ranked social status (Gobel et al., 2015). Other studies showed that when participants discriminate between emotional and neutral facial expressions, distinct fixation patterns emerge for each emotion. In particular, there is a focus on the lips for joyful faces and a focus on the eyes for sad faces (Schurgin et al., 2014). In parallel, it has very recently been shown that humans have idiosyncratic scanpaths while exploring faces (Kanan, Bseiso, Ray, Hsiao, & Cottrell, 2015) and that these scanning patterns are highly stable across time, representing a specific behavioural signature (Mehoudar, Arizpe, Baker, & Yovel, 2014). In the latter study, the authors asked the subjects to perform a face recognition task during three test sessions performed on three different days: Day 1, Day 3, and 18 months later. Their results show that individuals have very diverse scanning patterns. These patterns were not random but highly stable, even when examined 18 months later.

In this study, we aim to identify which factors drives this idiosyncratic behaviour. To quantify their relative contributions, we must overcome two major difficulties. First, in order to take into account as many factors as possible, we need an unprecedentedly large eye-tracking dataset. We have to move beyond the usual small-sized eye-tracking datasets with restricted participant profiles, the famous WEIRD (Western, Educated, Industrialized, Rich, and Democratic) population (Henrich, Heine, & Norenzayan, 2010). To do so, we increased our participant pool by recording and exploiting the largest and most diverse eye-tracking dataset that we are aware of (405 participants between 18 and 69 years old, 58 different nationalities) (Winkler & Subramanian, 2013). Second, we need to equip ourselves with eye-tracking data mining techniques that can identify and quantify eye-gaze patterns in the dataset. The vast majority of previous studies quantified observers' scanpaths through spatial distributions of eye positions averaged over time, deleting the - manifestly critical - temporal component of visual perception (Le Meur & Baccino, 2013). In contradistinction, we propose a new data-driven approach able to encapsulate the highly dynamic and individualistic dimension of a participant's gaze behaviour. We identified systemic differences that allow a classifier solely trained with eye tracking data to identify the gender of both the gazer and actor with very high accuracy.

Methods & Results

Experiment

Participants

We recorded the gaze of 459 visitors of the Science Museum of London, United Kingdom. We removed from the analysis the data of participants below 18 (n=8) as well as 46 other participants whose eye data exhibited some irregularities (loss of signal, obviously shifted positions). The analyses are performed on a final group of 405 participants (203 males, 202 females). Mean age of participants was M = 30.8, SD=11.5 (males: M = 32.3, SD = 12.3; females: M = 29.3, SD = 10.5). The experiment was approved by the UCL Research ethics committee and by the London Science Museum ethics board, and the methods were carried out in accordance with the approved guidelines. Signed informed consent was obtained from all participants.

Stimuli

Stimuli consisted in video clips of 8 different actors (4 females, 4 males, see Fig. 1). Each clip depicted the actor initially gazing towards the bottom of the screen for 500 ms, then gazing up at the participant for a variable amount of time (between 100 and 10,300 ms, in 300 ms increments across 35 clips) and finally gazing back at the bottom of the screen for 500 ms. Width x Height = 428 $\times$ 720 pixels (16.7 $\times$ 28.1 degrees of visual angle). Faces occupied most of the image, on average 280 $\times$ 420 pixels (10.9 $\times$ 16.4 degrees of visual angle). Average size of the eyes: 75 $\times$ 30 pixels (2.9 $\times$ 1.2 degrees of visual angle); nose: 80 $\times$ 90 pixels (3.1 $\times$ 3.5 degrees of visual angles); mouth: 115 $\times$ 35 pixels (4.5 $\times$ 1.4 degrees of visual angle). Frame rate = 30 Hz. Videos were shots with the same distance between the actors and the camera, in the same closed room with no window, in diffuse lighting conditions. Actors sat against a green background, and the point between their eyes (nasion) was aligned with the centre of the screen. Hence, the position of facial features slightly varied between actors due to individual morphological differences, but were largely overlapping between actors.

Apparatus

The experimental setup consisted of four computers: two for administering the personality questionnaire, and two dedicated to the eye-tracking experiment and actor face-rating questionnaire (see Procedure). Each setup consisted of a stimulus presentation PC (DELL precision T3500 & DELL precision T3610) hooked up to a 19" LCD monitor (both 1280 $\times$ 1024 pixels (49.9 $\times$ 39.9 degrees of visual angles) @ 60Hz) and an EyeLink 1000 kit (www.sr-research.com). Eye-tracking data was collected at 250 Hz. Participants sat at 57 cm from the monitor, their head stabilized with a chin rest, forehead rest, and headband. A protective opaque white screen encased the monitor and part of the participant's head in order to shield the participant from environmental distractions.

Procedure

The study took place at the Live Science Pod in the *Who Am I?* exhibition of the London Science Museum. The room had no window and the ambient luminance was very stable across the experiment. It consisted in 3 phases, for a total duration of approximately 15 minutes. Phase 1) A 10 items personality questionnaire based on the Big Five personality inventory (Rammstedt & John, 2007), collected on a dedicated set of computers. Each personality trait (extraversion, conscientiousness, neuroticism, openness and agreeableness) was assessed through 2 items. Item order was randomized across participants. Phase 2) The eye-tracking experiment. Experimental design features such as the task at hand or the initial gaze position have been shown to drastically impact on face exploration (Armann & Bülthoff, 2009; Arizpe, Kravitz, Yovel, & Baker, 2012). Here, they were standardized for every trial and participant. Participants were asked to freely look at 40 videos of one randomly selected actor. Prior to every trial, participants ~~had to look~~ at a black central fixation cross, presented on an average grey background. The fixation cross disappeared before the onset of the stimulus. On each trial one of 35 possible clips for that same actor was presented. Since there were 40 trials, some clips were presented twice. At the end of each trial participants were instructed to indicate via a mouse button press whether the amount of time the actor was engaged in direct gaze felt *uncomfortably short* or *uncomfortably long* with respect to what would feel appropriate in a *real* interaction. Each experiment was preceded by a calibration procedure, during which participants focused their gaze on one of nine separate targets in a 3 x 3 grid that occupied the entire display. A drift correction was carried out between each video, and a new calibration procedure was performed if the drift error was above 0.5 degree. Phase 3) An actor face rating questionnaire (Todorov, Said, Engell, & Oosterhof, 2008), where participants rated on a 1-7 scale the attractiveness, threat, dominance and trustworthiness of the actor they saw during the eye-tracking experiment.

Eye data processing We parsed the (x,y) eye position signal into fixations and saccades with a custom algorithm (Nyström & Holmqvist, 2010). This algorithm relies on an adaptive velocity threshold, making the event detection less sensitive to variations in noise level. Hence, detection thresholds varied between trials. Absolute thresholds are used to discard obvious outliers. We used the ones provided in (Nyström & Holmqvist, 2010), Table 2: max saccade velocity = 1000 degree/s; max saccade acceleration = 100,000 degree/s^2; min saccade duration = 10 ms, min fixation duration = 40 ms. Both position and pupil dilation data were further processed through a custom filtering algorithm that substituted signal losses with position / pupil data interpolated from data recorded prior and following the loss of signal. For every period of lost data (no size restriction) we performed a linear interpolation of the eye position (x and y coordinates) and variation in pupil size, using 100 ms of signal before and after loss. We did not set a limit on displacement between preceding and succeeding samples. We discarded all points that ~~felt~~ outside the screen. Pupil diameter was expressed on a trial-by-trial basis as a percentage change in diameter with respect to a baseline measure obtained in a 200 ms window preceding the start of each trial. Environment luminance was constant throughout the whole experiment duration. We post-processed the pupil signal to minimize artifacts caused by eye position. We removed the Pupil Foreshortening Effect artifact (Spring & Stiles, 1948; Jay, 1962) from the pupil data by implementing a correction technique, based on a geometric model that expresses the foreshortening of the pupil area as a function of the cosine of the angle between the eye-to-camera axis and the eye-to-stimulus axis (Hayes & Petrov, 2015). To estimate the variability of eye positions of a given observer, we used a dispersion metric (Coutrot, Guyader, Ionescu, & Caplier, 2012).

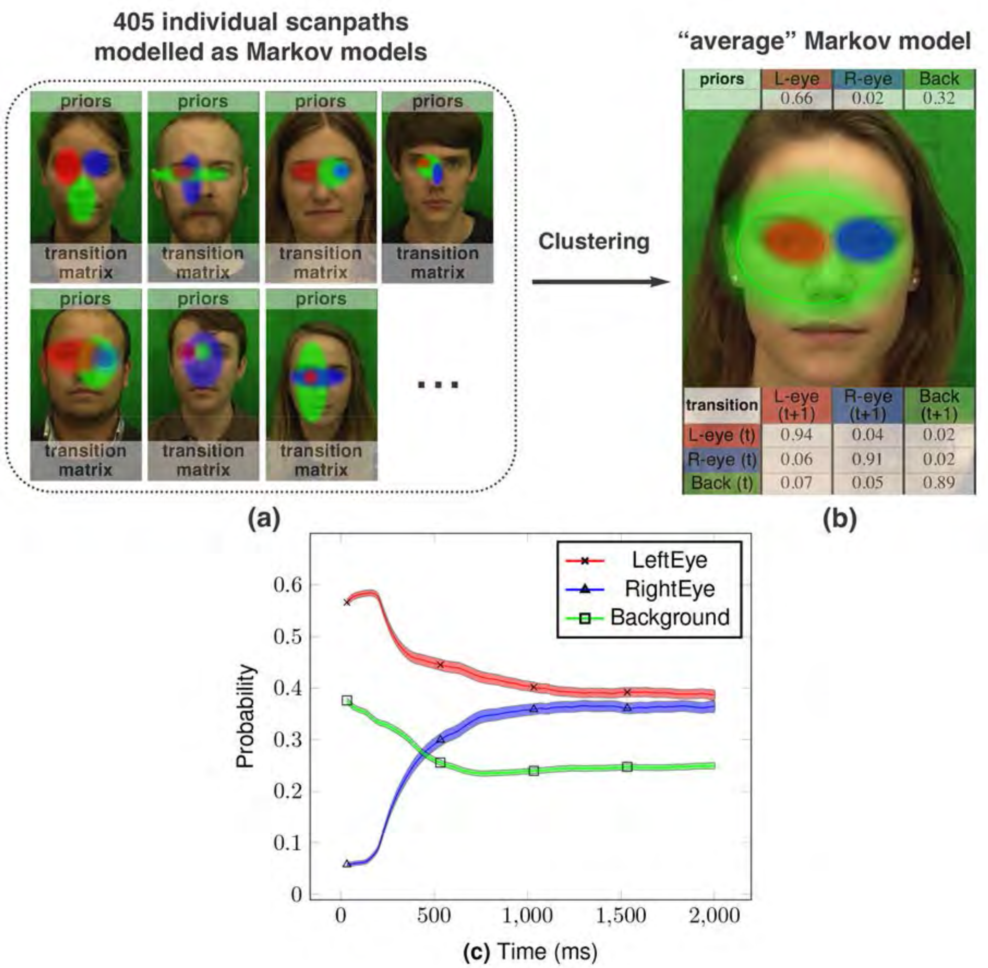

Figure 1: Scanpaths as Markov models. (a) Illustration of 7 out of the 405 individual scanpaths modelled as Markov models with 3 states. Each coloured area corresponds to a state, or region of interest. Transition matrices show the probabilities of staying and shifting. (b) Markov model averaged over the whole population with the VHEM algorithm. (c) Temporal evolution of the posterior probabilities of being in the states corresponding to fixating the left eye, right eye, and to the rest of the face. Error bars represent s.e.m. See also Supplementary Fig. S3.

If n eye positions were recorded from an observer ($\mathbf{p} = (x_i, y_i)_{i\in[1..n]}$), the eye position coordinates), the intra-observer dispersion D is defined as follows:

$$D(\mathbf{p}) = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{1}$$

The dispersion is the mean Euclidian distance between the eye positions of the same observers for a given clip. Small dispersion values reflect clustered eye positions.

Scanpath modelling

Hidden Markov Models (HMM)

To grasp the highly dynamic and individualistic components of gaze behaviour, we model participant's scanpaths using Hidden Markov Models (HMM) (Chuk, Chan, & Hsiao, 2014). This method acknowledges that visual exploration is a process that unfolds in a partic-

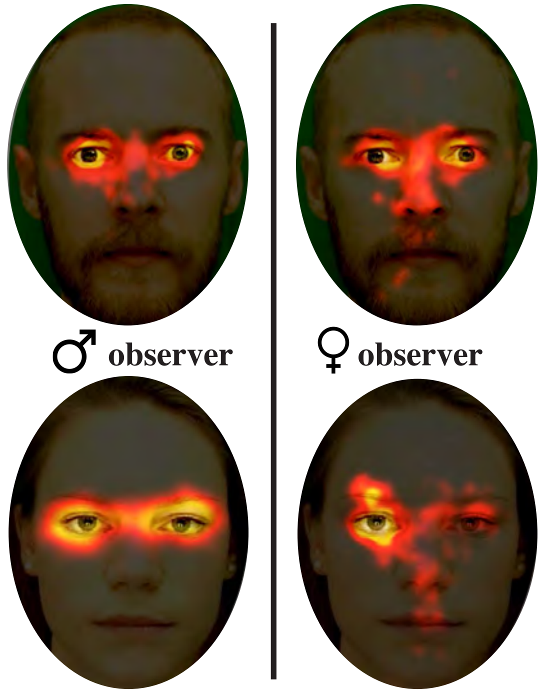

Figure 2: Heatmaps of eye positions of 4 representative individuals, with light colour indicating areas of intense focus. Left: male observers. Right: female observers. Top: male actors. Bottom: female actresses. One can see that females tend to explore faces more than males do, who stay more within the eye region. Females watching females have the strongest left-eye bias. See also Supplementary Fig. S4.

ular sequence over time, and cannot be summarised by mere average fixation counts or durations within predefined regions of interest. We first represent the average scanpath of each participant as a Markov model. A Markov model is a memory-less stochastic model used for representing probability distributions over sequences of observations, or states $(S_t)_{t\in[1..T]}$. It is composed of (1) priors (the probability distribution over the initial state) and (2) a state transition matrix, defining $P(S_t|S_{t-1})$, which encapsulates the probability of travel between states. States can denote processes (e.g. scanning / reading / decision) (Simola, Salojärvi, & Kojo, 2008), isolated targets (e.g. letter, line) (Haji-Abolhassani & Clark, 2013), but here, in the specific context of face exploration modelling, each state represents gaze falling on a region of interest (ROI) of the face. The distribution of eye positions (emission density) within each ROI is modelled as a 2D Gaussian distribution $N(m,\sigma)$, with m the centre and σ the covariance matrix (the generalization of variance to 2D) of the Gaussian. Modelling HMM states with Gaussians instead of more isolated targets allows to relax the point of interest, taking into account ~~phenomenons~~ such as the dissociation between the centre of gaze and the covert focus of attention, the imprecision of the human oculomotor system and of the eye-tracker. For more details on eye movement Bayesian modelling, we refer the reader to Boccignone's thorough introduction (Boccignone, 2015). All the HMM parameters (priors, between-state transition probabilities, mean and covariance of the Gaussians) are directly learnt from the eye data. They are obtained with the Baum-Welch algorithm, a special case of the Expectation-Maximization algorithm (Rabiner, 1989; Bishop, 2006). We set the prior means of the Gaussian emissions to be the actors' nasion, and the prior covariance matrix to be isotropic, with standard deviation of 200 pixels (roughly the same size as the facial features). We trained one model per participant using their eye positions subsampled at 30 Hz (1 value was drawn from the

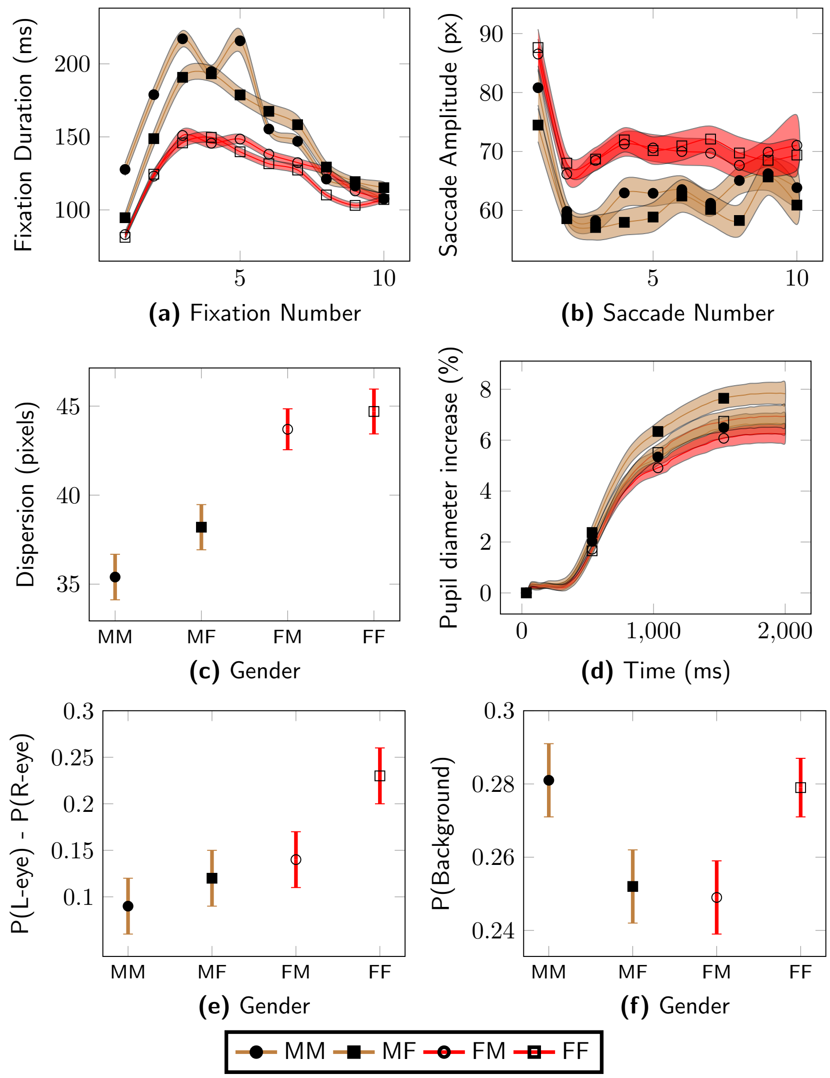

Figure 3: Gender differences in gaze behaviour. (a) Female observers make shorter fixations and (b) larger saccades. (c) Females are more scattered than male observers. (d) Increase in pupil diameter was expressed as a percentage change in diameter with respect to a baseline measure obtained in a 200 ms window preceding the start of each trial. Males watching females show a higher increase in pupil diameter than other pairings. (e) Females watching females (FF) have a stronger left-eye bias. (f) Males and females are both less likely to gaze at the eyes of a same-sex actor than of a different-sex actor. Error bars represent s.e.m.

raw signal every cyle). We chose this frequency for two reasons. 1) It conveniently matched the frame rate. 2) Training HMM with time-sampled eye positions instead of successive fixations allows ~~capturing~~ fixation duration information in the transition probabilities; 30 Hz is a good trade-off between complexity and scanpath information. To avoid local maximum problem, we repeated the training 3 times and retained the best performing model, ~~in term of~~ log-likelihood. We set the number of states to 3 (imposing the number of state is mandatory in order to cluster the Markov models) (Chuk et al., 2014). We also ~~tried~~ with 2 and 4 states, see Supplementary Fig. S1 and S2. With only 2 states, the model~~s~~ lacks ~~of~~ spatial resolution, and with K = 4 one state is often redundant~~, so~~ K = 3 is a good compromise. To cluster Markov models, we used the variational hierarchical EM algorithm (VHEM) for hidden Markov Models (Coviello, Chan, & Lanckriet, 2014). This algorithm clusters a given collection of Markov models into K groups of Markov models that are similar, according to their probability distributions, and characterizes each group by a 'cluster centre', i.e., a novel Markov Model that is representative for the group. For more details, the interested reader is referred to the original publication.

One advantage of this method is that it is totally data-driven; we have no prior hypothesis concerning the states parameters. To estimate the global eye movement pattern across all participants, we average the individual Markov models with the variational hierarchical EM algorithm (VHEM) for a Hidden Markov Model (Coviello et al., 2014). This algorithm clusters Markov models into N groups of similar models, and characterizes each group by a 'cluster centre', i.e. a model representative of that group. Here, we simply set $N = 1$. In Fig. 1a we show a representative set of 7 out of the 405 Markov Models with 3 states we trained with participants' eye positions sampled at 30 Hz. One can see the great variety of exploration strategies, with ROIs either distributed across the two eyes, the nose and the mouth, or shared between the two eyes, or even just focused on one eye. Yet, when Markov models are averaged over participants, the resulting centre model (Fig. 1b) displays a very clear spatial distribution: two narrow Gaussians on each eye and a broader one for the rest of the face. According to the priors, the probabilities at time zero, one is more likely to begin exploring the face from the left eye (66 %), than from the background (32 %) or from the right eye (only 2 %). Moreover, the transition matrix states that when one is looking at the left eye, one is more likely to stay in this state (94 %, with a 30 Hz sampling rate) than when in the right eye region (91 %) or in the rest of the face (back, 89 %). These values are backed-up by the temporal evolution of the posterior probability of each state (Fig. 1c). We clearly see a very strong left-eye bias during the first 250 milliseconds. This bias persists throughout but decreases over time.

Clusters of gaze behaviour

To test whether this general exploration strategy accounts for the whole population or if there are clusters of observers with different gaze behaviours, we again use the VHEM algorithm. We set $N = 2$ (other values have been tested, but did not lead to significant differences between groups) and obtain two distinct patterns (see Supplementary Fig. S3). To identify the variables leading to these different gaze strategies, we use Multivariate Analysis of Variance (MANOVA). MANOVA seeks to determine whether multiple levels of independent variables, on their own or in combination, have an effect on the affiliation of participants to one group or the other. We tested 3 groups of independent variables: personality traits, face ratings, and basic characteristics (nationality, age, gender of the observer (GO), gender of the corresponding actor (GA) and identity of the actor). Only the basic characteristics group led to a significant separation between the two clusters ($F_{1,404} = 7.5, p = 0.01$), with observer and actor gender having the highest coefficient absolute value (see Supplementary Table S1). Face ratings and personality traits failed to account for the separation between the two clusters (resp. $F_{1,404} = 1.4, p = 0.58$ and $F_{1,404} = 2.0, p = 0.55$). We also ran a MANOVA with the three groups of independent variables combined and obtained similar results. We performed the same analysis for HMMs with 2 and 4 states and obtained similar results, see Supplementary Table S1 for basic characteristics MANOVA coefficients. Hence, GA and GO are the most efficient variables to use to separate exploration strategies into two subgroups. In the following, we closely characterise these gender-induced gaze differences.

Gender differences

To characterise gender differences in face exploration, we split our dataset into 4 groups: male observers watching male actors (MM,

$n = 119$), male observers watching female actors (MF, $n = 84$), female observers watching male actors (FM, $n = 106$) and female observers watching female actors (FF, $n = 96$). Fig. 2 displays the eye position heatmaps of 4 individuals, each of them being characteristic of one of these groups. We first compared the simplest eye movement parameters between these groups: saccade amplitudes, fixation durations and intra-participant dispersion (i.e. eye position variance of a participant within a trial). In the following, the statistical significance of the effects of actor and observer gender has been evaluated via one-way ANOVAs. Pair-wise comparisons have been explored via Tukey's post-hoc comparisons. We find that for male observers, fixation durations are longer (Fig. 3a, $F_{3,401} = 10.1, p < 0.001$), saccade amplitudes shorter (Fig. 3b, $F_{3,401} = 8.5, p < 0.001$), and dispersion smaller (Fig. 3c, $F_{3,401} = 12.9, p < 0.001$) than for female observers. Actor gender does not influence these results. Note that these results are mutually consistent: shorter saccades and longer fixations logically lead to lower dispersion. Gender also impacts on the increase of pupil diameter: this value is greater in MF than in any other group (Fig. 3d, $F_{3,401} = 2.8, p = 0.03$), consistent with the Belladonna effect (Tombs & Silverman, 2004; Rieger et al., 2015). We used the VHEM algorithm to obtain the centre Markov model of these 4 groups ($N = 1$ within each group). In all 4 groups, the spatial distribution of states is similar to the one depicted in Fig. 1b. We computed the posterior probabilities of each state, as we did for the whole population (Supplementary Fig. S4). We find that during the first second of exploration, the left-eye bias is stronger in FF than in MM (Fig. 3e, $F_{3,401} = 3.1, p = 0.02$), with no difference between the other comparisons. We also show that one is less likely to gaze at the eyes of a same-sex actor than of a different-sex actor (Fig. 3f, $F_{3,401} = 2.9, p = 0.03$).

Gaze-based classification

These patterns appear systematic and rich enough to differentiate both actor and observer gender. To test this we gathered all the gender differences in gaze behaviour mentioned so far to train a classifier to be able to predict the gender of a given observer and/or the gender of the observed face (Fig. 4). We started from a set of 15 variables $(D_i)_{i\in[1..N]}$, with N the total number of participants. Vector D_i gathers for participant i the following information. Markov model parameters (spatial (x,y) coordinates of the states ranked by decreasing posterior probability value, posterior probability of the left eye, right eye and rest of the face averaged over the first second of exploration), mean intra-participant dispersion, mean saccade amplitude, mean fixation duration, mean pupil diameter, peak pupil diameter and latency to peak. We then reduced the dimensionality of this set of variables $(D_i)_{i\in[1..N]}$ by applying a MANOVA. We applied two different MANOVAs: one to optimize the separation between 2 classes (M observers vs. F observers), the other to optimize the separation between 4 classes (MM vs. MF vs. FM vs. FF). The eigenvector coefficients corresponding to each variable are available in Supplementary Table S2 and Supplementary Table S3. To infer the gender of an observer j, we used Quadratic Discriminant Analysis (QDA). We followed a Leave-One-Out approach: at each iteration, one participant was taken out ~~out for~~ test, and the classifier trained with all the others. For each participant j, we trained a QDA-classifier with $(EV_i)^{i\neq j}_{i\in[1..N]}$, EV representing the first two eigenvectors from the first MANOVA. To infer the gender of both observer j and of the corresponding actor, we followed the same approach with the first two eigenvectors of the second MANOVA. Both classifiers perform highly above chance level (two-sided binomial test, $p < 0.001$). Such a classifier is able to correctly guess the gender of the observer 73.4% of the time (2 classes, chance level = 50%). It correctly guesses the gender of both the observer and of the corresponding face 51.2% of the time (4 classes, chance level = 25%). We obtained

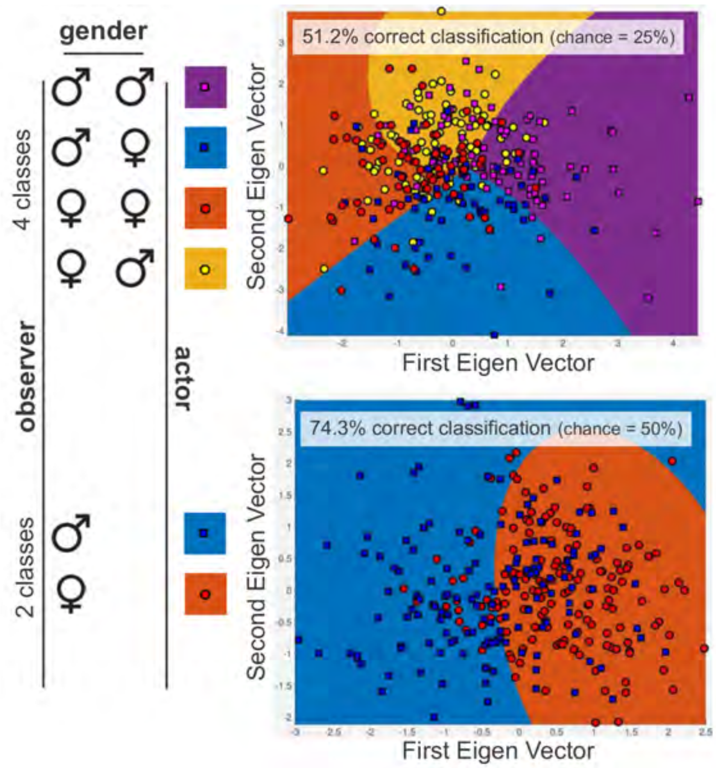

Figure 4: Gaze-Based Gender Classification. Solely relying on eye data, a classifier based on dimensionality reduction and Quadratic Discriminant Analysis achieves 51.2% of correct classification with 4 classes (MM, MF, FM FF, chance level = 25%), and 74.3% with 2 classes (M or F observers, chance level = 50%). Both classifiers perform highly above chance (one-sided binomial test, $p < 0.001$). The initial dataset consists of: spatial coordinates of the states ranked by posterior probability value, associated posterior probability averaged over the first second of exploration, intra-participant dispersion, mean saccade amplitude, mean fixation duration, mean dispersion, mean pupil diameter, peak pupil diameter and latency to peak. See also Supplementary Table S2 and Supplementary Table S3.

similar results with Linear Discriminant Analysis (LDA): 54.3% correct classification with 4 classes, 73.4 % correct classification with 2 classes.

Discussion

Understanding the precise nature of face perception is challenging as the face comprises a high-dimensional, dynamic information space (Jack & Schyns, 2015). In this study, we use novel data-mining methods that encapsulate the highly dynamic and individualistic spatio-temporal nature of gaze. Although a few previous studies have used Markov-based analysis with eye-tracking data to identify fixations and saccades (Salvucci & Goldberg, 2000), to infer observers' tasks (Simola et al., 2008; Haji-Abolhassani & Clark, 2014), or to build visual saliency models (Zhong, Zhao, Zou, Wang, & Wang, 2014), only a small number of recent studies have applied these techniques to face exploration (Chuk et al., 2014; Kanan et al., 2015). This approach is particularly powerful as faces feature very clear and stable regions of interest (eyes, mouth, nose), allowing meaningful comparisons of Markov model states across stimuli and observers. Here, for the first time, we propose to jointly use Bayesian (Markov model clustering) and Frequentist (MANOVA) inferences to assess the influence of a large set of variables on face exploration strategies. We tested variables related to observers' psychological profile (personality traits), how they perceived actors' face (face ratings), as well as basic demographic such as age, nationality or gender. We found both the gender of the observer and of the actor to be the most efficient variables to separate the different recorded exploration strategies into two homogeneous subgroups. This outcome cannot be explained by differences between stimuli since observers' gender is balanced between stimuli, as shown in Supplementary Fig. S5. This is backed-up by the very low MANOVA coefficient associated with the identity of the actor (7.9e-3), see Supplementary Table S1. Our model-based results are supplemented with more classic~~al~~ eye movement parameters such as fixation duration, saccade amplitude, intra-observer dispersion, and pupilometry. In the following, we discuss the different strategies followed by different gender groups as well as their implication is perception bias.

Males are less exploratory than females

We present three complementary metrics indicating that female observers explore the face they are presented with more, ~~regardless~~ of the gender of the actor: females make shorter fixations, larger saccades, and their eye positions are more scattered over the actor face. Previous studies have reported the same pattern, even with very different stimuli or experimental designs (Vassallo, Cooper, & Douglas, 2009; Shen & Itti, 2012; Mercer Moss, Baddeley, & Canagarajah, 2012). For instance, in (Shen & Itti, 2012), eye movements were recorded while participants watched and listened to different speakers in various outdoor settings. The authors showed that women saccade more often away from the face of the speaker, especially to ~~his/her~~ body. This difference in gaze behaviour has been linked to the greater accuracy of women in the decoding of nonverbal cues (Hall, 1984; McClure, 2000; Hall & Matsumoto, 2004; Schmid, Mast, Bombari, & Mast, 2011). Actively looking for nonverbal cues distributed in many different parts of the face (Vatikiotis-Bateson & Kuratate, 2012) and body of the speakers, especially their hands (Krauss, Chen, & Chawla, 1996), would increase ~~female~~ gaze dispersion. However, this hypothesis is undermined by *Reading the Mind in the Eyes Test* developed in (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), where participants are asked to associate photographs of pairs of eyes with an adjective (e.g., playful, comforting, irritating, bored). Indeed, females are better than males at this task, even though only the eye region is available (Kirkland, Peterson, Baker, Miller, & Pulos, 2013). Furthermore, males have been shown to be less exploratory than females even when exploring visual scenes without any face such as landscapes or art stimuli (Mercer Moss et al., 2012).

Females looking at females have stronger left-side bias

We provide some insights into the time course of the global preference for the eye on the left side. We report a very strong left-eye bias during the first 250 milliseconds of exploration, ~~persisting~~ ~~throughout but decreasing~~ over time.The left-eye bias is a well documented, face specific characteristic. This bias is very strong when exploring upright faces, ~~weakens with~~ inverted faces, and disappears with non-face stimuli, whether they are symmetric (vases, fractals) or not (landscapes) (Mertens, Siegmund, & Gruesser, 1993; Leonards & Scott-Samuel, 2005). It is often associated with chimeric faces: faces composed of two left halves are often judged to be more similar to the original face than faces composed of two right halves. Other studies have reported than when the left and the right-hand side of a face are different, observers tend to base their responses on the information contained in the left side. This includes face recognition (Brady, Campbell, & Flaherty, 2005), gender identification (Butler et al., 2005), facial attractiveness, expression and age (Burt & Perrett, 1997). The factors determining this lateralization remain unclear. Some obvious potential determining factors have been excluded, such as observers' eye or hand dominance (Leonards & Scott-Samuel, 2005). Other factors has been shown to interact with the left-side bias, such as scanning habits (left bias is weakened for native readers of right to left) (Megreya & Havard, 2011), and start position (bias toward facial features furthest from the start position) (Arizpe et al., 2012; Arizpe, Walsh, & Baker, 2015). The most common explanation found in the literature involves the right hemisphere ~~domination~~ for face processing (Kanwisher, McDermott, & Chun, 1997; Yovel, Tambini, & Brandman, 2008). Although valid when fixation location is fixed at the centre of the face, it may seem counterintuitive when participants are free to move their eyes (Everdell, Marsh, Yurick, Munhall, & Paré, 2007; Hsiao & Cottrell, 2008; Guo, Meints, Hall, Hall, & Mills, 2009). Indeed, looking at the left side of the stimulus places the majority of the actor's face ~~information~~ in the observer's right visual field i.e. their left hemisphere. An interpretation proposed in (Butler et al., 2005) is that starting from a central position - either because of an initial central fixation cross or because of the well-known centre bias (Tseng, Carmi, Cameron, Munoz, & Itti, 2009) - the left-hand side of the face activates right hemisphere face functions, making the latter initially more salient than its right counterpart. This interpretation is also supported by the fact that during a face identification task, the eye on the left side of the image becomes diagnostic before the one on the right (Vinette, Gosselin, & Schyns, 2004; Rousselet, Ince, Rijsbergen, & Schyns, 2014). This explains why the left-side bias is so strong during the very first moments of face exploration, but why does it persists over time? Different authors have reported a preferred landing position between the right eye (left visual hemifield) and the ~~left~~ of the nose during face recognition (Hsiao & Cottrell, 2008; Peterson & Eckstein, 2013). This places a region of dense information - the left eye and brow - within the foveal region, slightly displaced to the left visual hemifield, hence again activating right hemisphere face processing functions. An alternative hypothesis is that the left side bias could be linked to the prevalence of right eye dominance in humans. When engaged in mutual gaze, the dominant eye ~~provides~~ the best cue to gaze direction through small vergence cues. Since the majority of the population is right-eye dominant (Coren, 1993), humans might prefer looking at the right eye (i.e. at the left side of the face) as it provides a clearer signal of mutual gaze. Here, we found the left-side bias stronger in females looking at other females. This strengthening, coupled with the fact that the perception of facial information is biased towards the left side of face images complements an earlier study reporting that females are better at recognizing other female faces, whereas there are no gender differences with regard to male faces (Rehnman & Herlitz, 2006). On the other hand, it is inconsistent with another study reporting that when looking at faces expressing a variety of emotions, men show an asymmetric functioning of visual cortex, whereas women have a more bilateral functioning (Proverbio, Brignone, Matarazzo, Del Zotto, & Zani, 2006). Further investigation is needed to disentangle the interaction between the gender of the observer and of the face observed in the activation of the right hemisphere face processing functions.

Limitations of this study

The authors would like to make clear that this study does not demonstrate that gender is the variable that most influences gaze patterns during face exploration *in general*. Many aspects of the experimental design might have influenced the results presented in this paper. The actors we used were all Caucasian between 20 and 40 years old with a neutral expression and did not speak, all factors that could have influenced observers' strategies (Wheeler et al., 2011; Schurgin et al., 2014; Coutrot & Guyader, 2014). Even the initial gaze position has been shown to have a significant impact on the following scanpaths (Arizpe et al., 2012, 2015). In particular, the task given to the participants - rating the level of comfort they felt with the actor's duration of direct gaze - ~~has~~ certainly biase~~d~~ participants' attention toward actors' eyes. One of the first eye-tracking experiment in history suggested that gaze patterns are strongly modulated by different task demands (Yarbus, 1965). This result has since been replicated and extended: more recent studies showed that the task at hand can even be inferred using gaze-based classifiers (Borji & Itti, 2014; Haji-Abolhassani & Clark, 2014; Kanan et al., 2015; Boisvert & Bruce, 2016). Here, gender appears to be the variable ~~inducing~~ the strongest differences between participants. But one could legitimately hypothesize that if the task had been to determine the emotion displayed by the actors' face, the culture of the observer could have play a more important role, as it has been shown that the way we perceive facial expression is not universal (Jack, Blais, Scheepers, Schyns, & Caldara, 2009). Considering the above, ~~we believe that~~ the key message of this paper is that ~~given a set of stimuli and an experimental design, with all their inherent idiosyncrasies,~~ our method allows capturing systematic differences between groups of observers *in a data-driven fashion*.

Conclusion

Using the biggest and most diverse eye-tracking database recorded, we show that the way people look at faces contains systematic variations that are diagnostic of the gender of the observers and of the face they observe. These results have ~~strong~~ implications in every field using gaze-based models. For instance, quantifying the nature of face processing and joint attention is critical to the understanding and the diagnosis of disorders such as Schizophrenia, Autism or Attention Deficit Hyperactivity Disorder (Freeth, Foulsham, & Kingstone, 2013; Wang et al., 2015). Tailoring these gaze-based models to a masculine or feminine population could lead to significant enhancements, particularly when a substantial sex ratio difference exists (e.g. Autism). Going even further, one can speculate that different stimuli could elicit different systematic patterns diagnostic of other observers' characteristics, such as their state of health or level of cognitive development (Tseng et al., 2013; Wass & Smith, 2014). Given the ubiquitous nature of eye movements, being able to deduce such fundamental characteristics about a person, without the need for self report, would have tremendous impact across a broad range of fields.

Competing interests We have no competing interests.

Funding Sources A.C. has been funded by the Engineering and Physical Sciences Research Council (Grant No. EP/I017909/1). N.B., C.H., I.M. and A.J. have been funded by a Leverhulme trust grant (RPG-2013-218).

Data accessibility The datasets supporting this article will be uploaded on Dryad.

Appendix

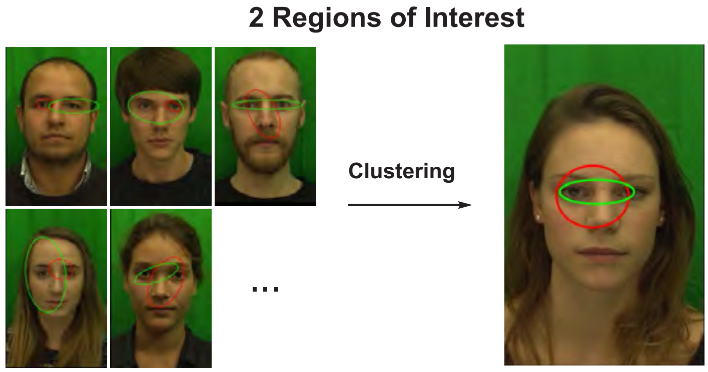

Figure S 1: Left: illustration of 5 out of the 405 individual scanpaths modelled as Markov models with 2 states. Each coloured area corresponds to a state, or region of interest. Right: Markov model averaged over the whole population with the VHEM algorithm.

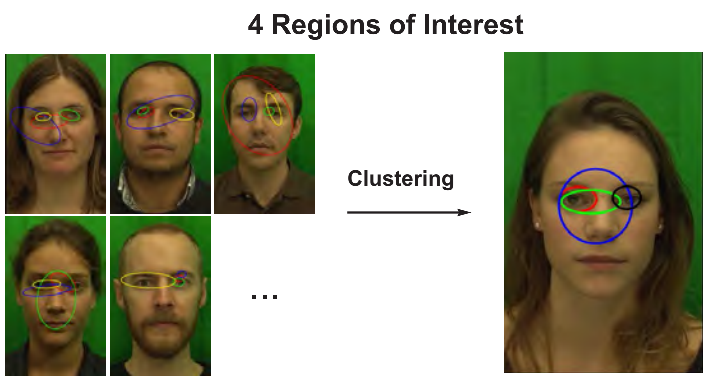

Figure S 2: Left: illustration of 5 out of the 405 individual scanpaths modelled as Markov models with 4 states. Each coloured area corresponds to a state, or region of interest. Right: Markov model averaged over the whole population with the VHEM algorithm.

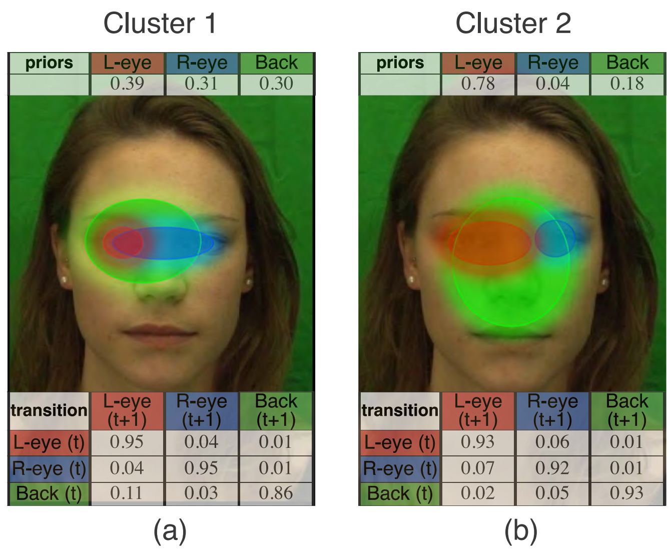

Figure S 3: Two clusters of gaze behaviour. Centres of the two clusters of Markov models with 3 states obtained via the VHEM algorithm. Cluster 1 features narrow Gaussians centred on the eye region. Its priors are balanced between the 3 states. The transition probabilities are stronger for the eyes than for the rest of the face. Cluster 2 features broader states, with priors favouring the left eye over the right eye. The transition probabilities are balanced between the three states.

(a) - HMMs with 2 Regions of Interest

Independant Variables	Observer Gender	Observer Age	Observer Nationality	Actor Number	Actor Gender
First Eigen Vector	8.1e-1	8.1e-2	4.6e-2	- 8.1e-2	1.4e-1

(b) - HMMs with 3 Regions of Interest

Independant Variables	Observer Gender	Observer Age	Observer Nationality	Actor Number	Actor Gender
First Eigen Vector	4.5e-1	5.6e-2	- 2.1e-1	7.9e-3	2.8e-1

(c) - HMMs with 4 Regions of Interest

Independant Variables	Observer Gender	Observer Age	Observer Nationality	Actor Number	Actor Gender
First Eigen Vector	- 6.1e-1	1.0e-2	2.6e-2	1.5e-1	1.8

Table S 1: Coefficients of the first eigenvector of the MANOVA separating the two clusters of HMMs computed via the VHEM algorithm. We performed this analysis for HMMs with (a) 2 states, (b) 3 states, and (c) 4 states. Categorical variable: cluster 1 or 2. Independent variables: basic characteristics group. This group is the only one that led to a significant separation between the two clusters of Markov models. In the three analyses, the highest coefficient absolute values are the ones of Observer Gender and Actor Gender.

state coord. ranked by posterior probabilities						posterior probabilities			classic parameters			pupillometry		
x_1	y_1	x_2	y_2	x_3	y_3	$p_{lefteye}$	$p_{righteye}$	p_{back}	sacc amp	fix dur	disp	p.mean	p.peak	p.lat
First Eigen vector														
7.8e-3	-3.5e-3	5.4e-3	-3.0e-3	1.4e-2	-3.9e-3	2.8	2.0	2.7	-1.5e-4	-6.9e-3	1.4e-2	1.5e-4	1.7e-3	-9.2e-3
Second Eigen vector														
2.7e-3	9.9e-3	-6.7e-4	1.0e-2	-2.3e-3	1.0e-2	13.8	13.8	13.0	2.7e-2	-1.8e-3	-2.9e-2	-2.8e-3	-3.4e-3	1.7e-3

Table S 2: Coefficients of the first and second eigenvectors of the MANOVA separating the eye data into 4 classes: MM, MF, FM and FF. Categorical variables: gender of the observer and of the actor. Independent variables: state coordinates ranked by decreasing mean posterior probabilities, mean posterior probabilities of the left eye, right eye and background, mean saccade amplitude, fixation duration and intra-observer dispersion, mean pupil diameter, maximum pupil diameter and latency to maximum pupil diameter. The ratio of the between-group variance to the within-group variance for the first Eigen vector is 0.39; 0.23 for the second.

state coord. ranked by posterior probabilities						posterior probabilities			classic parameters			pupillometry		
x_1	y_1	x_2	y_2	x_3	y_3	$p_{lefteye}$	$p_{righteye}$	p_{back}	sacc amp	fix dur	disp	p.mean	p.peak	p.lat
First Eigen vector														
8.5e-3	5.8e-4	4.7e-3	1.5e-3	9.9e-3	2.6e-3	8.9	8.2	8.2	1.3e-2	-6.7e-3	-2.3e-4	1.0e-3	-2.4e-4	-7.0e-3
Second Eigen vector														
-5.8e-4	-2.7e-3	-1.1e-2	-1.0e-2	-2.2e-7	-4.7e-2	-5.7e-1	-4.2e-1	-9.8e-1	3.0e-3	-1.8e-3	-1.1e-2	-7.3e-4	-1.4e-3	-5.6e-4

Table S 3: Coefficients of the first and second eigenvectors of the MANOVA separating the eye data into 2 classes: Male and Female observer. Categorical variables: gender of the observer. Independent variables: state coordinates ranked by decreasing mean posterior probabilities, mean posterior probabilities of the left eye, right eye and background, mean saccade amplitude, fixation duration and intra-observer dispersion, mean pupil diameter, maximum pupil diameter and latency to maximum pupil diameter. The ratio of the between-group variance to the within-group variance for the first Eigen vector is 0.32; 3e-16 for the second.

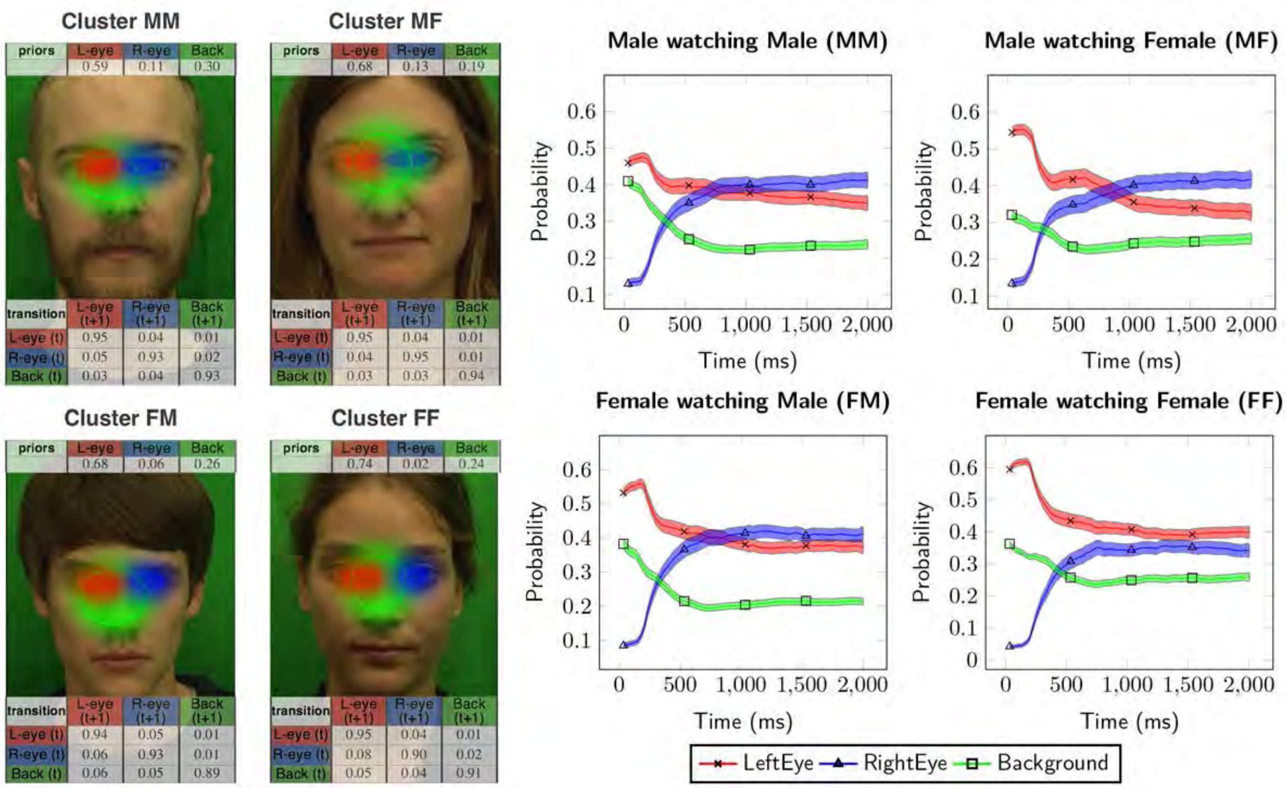

Figure S 4: Clusters of Markov model by observer and actor gender. Left: Markov models belonging to the MM, MF, FM or FF group clustered via VHEM algorithm. Right: Corresponding posterior probabilities, for the three possible states (left eye, right eye and rest of the face). Error bars represent s.e.m.

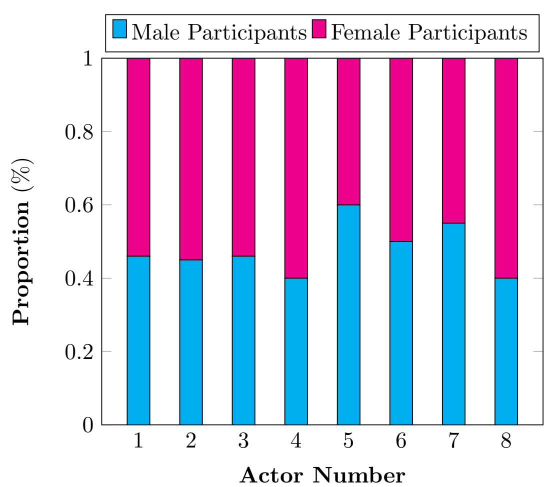

Figure S 5: For each actor, proportion of male and female participants. Actor 2, 3, 6 and 7 are male, the other are female. The analyses are performed on 405 participants (203 males, 202 females).

References

Arizpe, J. M., 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.

Arizpe, J. M., Walsh, V., & Baker, C. I. (2015). Characteristic visuomotor influences on eye-movement patterns to faces and other high level stimuli. *frontiers in Psychology*, *6*, 1–14.

Armann, R., & Bülthoff, I. (2009). Gaze behavior in face comparison: The roles of sex, task, and symmetry. *Attention, Perception, & Psychophysics*, *71*(5), 1107–1126.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The 'reading the mind in the eyes' test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, *42*(2), 241–251.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. New York, USA: Springer.

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture Shapes How We Look at Faces. *PLoS ONE*, *3*(8), e3022.

Boccignone, G. (2015). Advanced statistical methods for eye movement analysis and modeling: a gentle introduction. *arXiv*, *1506.07194*.

Boisvert, J. F. G., & Bruce, N. D. B. (2016). Predicting task from eye movements: On the importance of spatial distribution, dynamics, and image features. *Neurocomputing*, 1–36.

Borji, A., & Itti, L. (2014). Defending Yarbus: Eye movements reveal observers' task. *Journal of Vision*, *14*(3), 1–22.

Borji, A., Lennartz, A., & Pomplun, M. (2015). What do eyes reveal about the mind? Algorithmic inference of search targets from fixations . *Neurocomputing*, *149*(PB), 788–799.

Brady, N., Campbell, M., & Flaherty, M. (2005). Perceptual asymmetries are preserved in memory for highly familiar faces of self and friend. *Brain and cognition*, *58*(3), 334–342.

Burt, M. D., & Perrett, D. I. (1997). Perceptual Asymmetries in Judgements of Facial Attractiveness, Age, Gender, Speech and Expression. *Neuropsychologia*, *35*, 685–693.

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.

Chuk, T., Chan, A. B., & Hsiao, J. H. (2014). Understanding eye movements in face recognition using hidden markov models. *Journal of vision*, *14*(11), 8.

Coren, S. (1993). The lateral preference inventory for measurement of handedness, footedness, eyedness, and earedness: Norms for young adults. *Bulletin of the Psychonomic Society*, *31*(1), 1–3.

Coutrot, A., & Guyader, N. (2014). How saliency, faces, and sound influence gaze in dynamic social scenes. *Journal of Vision*, *14*(8), 1–17.

Coutrot, A., Guyader, N., Ionescu, G., & Caplier, A. (2012). Influence of soundtrack on eye movements during video exploration.

Journal of Eye Movement Research, *5*(4), 1–10.

Coviello, E., Chan, A. B., & Lanckriet, G. R. G. (2014). Clustering Hidden Markov Models with Variational HEM. *Journal of Machine Learning Research*, *15*, 697–747.

Eisenbarth, H., & Alpers, G. W. (2011). Happy Mouth and Sad Eyes: Scanning Emotional Facial Expressions. *Emotion*, *11*(4), 860–865.

Everdell, I. T., Marsh, H., Yurick, M. D., Munhall, K. G., & Paré, M. (2007). Gaze behaviour in audiovisual speech perception: Asymmetrical distribution of face-directed fixations. *Perception*, *36*, 1535–1545.

Foulsham, T., Cheng, J. T., Tracy, J. L., Henrich, J., & Kingstone, A. (2010). Gaze allocation in a dynamic situation: Effects of social status and speaking. *Cognition*, *117*(3), 319–331.

Freeth, M., Foulsham, T., & Kingstone, A. (2013). What Affects Social Attention? Social Presence, Eye Contact and Autistic Traits. *PLoS ONE*, *8*(1), e53286.

Gobel, M. S., Kim, H. S., & Richardson, D. C. (2015). The dual function of social gaze. *Cognition*, *136*(C), 359–364.

Guo, K., Meints, K., Hall, C., Hall, S., & Mills, D. (2009). Left gaze bias in humans, rhesus monkeys and domestic dogs. *Animal Cognition*, *12*, 409–418.

Haji-Abolhassani, A., & Clark, J. J. (2013). A computational model for task inference in visual search. *Journal of Vision*, *13*(3), 1–24.

Haji-Abolhassani, A., & Clark, J. J. (2014). An inverse Yarbus process: Predicting observers' task from eye movement patterns. *Vision Research*, *103*, 127–142.

Hall, J. A. (1984). *Nonverbal sex differences: Communication accuracy and expressive style*. Baltimore, Maryland, USA: The Johns Hopkins University Press.

Hall, J. A., & Matsumoto, D. (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion*, *4*(2), 201.

Hayes, T. R., & Petrov, A. A. (2015). Mapping and Correcting the Influence of Gaze Position on Pupil Size Measurements. *Behavior Research Methods*, *First online*, 1–16.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2-3), 61–135.

Hsiao, J. Hui-wen, & Cottrell, G. W. (2008). Two fixations suffice in face recognition. *Psychological Science*, *19*(10), 998–1006.

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.

Jack, R. E., & Schyns, P. G. (2015). The Human Face as a Dynamic Tool for Social Communication. *Current Biology*, *25*(14), R621–R634.

Jay, B. (1962). The effective pupillary area at varying perimetric angles. *Vision Research*, *1*(5-6), 418–424.

Kanan, C., Bseiso, D. N. F., Ray, N. A., Hsiao, J. Hui-wen, & Cottrell, G. W. (2015). Humans Have Idiosyncratic and Task-specific Scanpaths for Judging Faces. *Vision Research*, *108*, 67–76.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *The Journal of neuroscience*, *17*(11), 4302–4311.

Kirkland, R. A., Peterson, E., Baker, C. A., Miller, S., & Pulos, S. (2013). Meta-analysis Reveals Adult Female Superiority in" Reading the Mind in the Eyes Test". *North American Journal of Psychology*, *15*(1), 121–146.

Krauss, R. M., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? *Advances in experimental social psychology*, *28*, 389–450.

Le Meur, O., & Baccino, T. (2013). Methods for comparing scanpaths and saliency maps: strengths and weaknesses. *Behavior Research Methods*, *45*(1), 251–266.

Leonards, U., & Scott-Samuel, N. E. (2005). Idiosyncratic initiation of saccadic face exploration in humans. *Vision Research*, *45*, 2677–2684.

McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological bulletin*, *126*(3), 424.

Megreya, A. M., & Havard, C. (2011). Left face matching bias: Right hemisphere dominance or scanning habits? *Laterality: Asymmetries of Body, Brain and Cognition*, *16*(1), 75–92.

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), 1–11.

Mercer Moss, F. J., Baddeley, R., & Canagarajah, N. (2012). Eye Movements to Natural Images as a Function of Sex and Personality. *PLoS ONE*, *7*(11), 1–9.

Mertens, I., Siegmund, H., & Gruesser, O.-J. (1993). Perceptual asymmetries are preserved in memory for highly familiar faces of self and friend. *Neuropsychologia*, *31*, 989–998.

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods*, *42*(1), 188–204.

Perlman, S. B., Morris, J. P., Vander Wyk, B. C., Green, S. R., Doyle, J. L., & Pelphrey, K. A. (2009). Individual Differences in Personality Predict How People Look at Faces. *PLoS ONE*, *4*(6), e5952.

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.

Proverbio, A. M., Brignone, V., Matarazzo, S., Del Zotto, M., & Zani, A. (2006). Gender differences in hemispheric asymmetry for face processing. *BMC Neuroscience*, *7*(44).

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257–286.

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*, 203–212.

Rehnman, J., & Herlitz, A. (2006). Higher face recognition ability in girls: Magnified by own-sex and own-ethnicity bias. *Memory*, *14*(3), 289–296.

Rieger, G., Cash, B. M., Merrill, S. M., Jones-Rounds, J., Dharmavaram, S. M., & Savin-Williams, R. C. (2015). Sexual arousal: The correspondence of eyes and genitals. *Biological Psychology*, *104*, 56–64.

Rousselet, G. A., Ince, R. A. A., Rijsbergen, N. J. v., & Schyns, P. G. (2014). Eye coding mechanisms in early human face event-related potentials. *Journal of Vision*, *14*(13), 1–24.

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying Fixations and Saccades in Eye-Tracking Protocols. In *Proceedings of the symposium on eye tracking research applications* (pp. 71–78).

Schmid, P. C., Mast, M. S., Bombari, D., & Mast, F. W. (2011). Gender effects in information processing on a nonverbal decoding task. *Sex Roles*, *65*(1-2), 102–107.

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), 1–16.

Shen, J., & Itti, L. (2012). Top-down influences on visual attention during listening are modulated by observer sex. *Vision Research*, *65*(C), 62–76.

Simola, J., Salojärvi, J., & Kojo, I. (2008). Using hidden Markov model to uncover processing states from eye movements in information search tasks. *Cognitive Systems Research*, *9*(4), 237–251.

Spring, K., & Stiles, W. (1948). Apparent shape and size of the pupil viewed obliquely. *British Journal of Ophthalmology*, *32*, 347–354.

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, *12*, 455–460.

Tombs, S., & Silverman, I. (2004). Pupillometry - A sexual selection approach . *Evolution and Human Behavior*, *25*(4), 221–228.

Tseng, P.-H., Cameron, I. G. M., Pari, G., Reynolds, J. N., Munoz, D. P., & Itti, L. (2013). High-throughput classification of clinical populations from natural viewing eye movements. *Journal of Neurology*, *260*, 275–284.

Tseng, P.-H., Carmi, R., Cameron, I. G. M., Munoz, D. P., & Itti, L. (2009). Quantifying center bias of observers in free viewing of dynamic natural scenes. *Journal of Vision*, *9*(7), 1–16.

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*(3), 1–10.

Vatikiotis-Bateson, E., Eigsti, I.-M., Yano, S., & Munhall, K. G. (1998). Eye movement of perceivers during audiovisual speech perception. *Perception & Psychophysics*, *60*(6), 926–940.

Vatikiotis-Bateson, E., & Kuratate, T. (2012). Overview of audiovisual speech processing. *Journal of Acoustical Science and Technology*, *33*(3), 135–141.

Vinette, C., Gosselin, F., & Schyns, P. G. (2004). Spatio-temporal dynamics of face recognition in a flash: it's in the eyes. *Cognitive Science*, *28*, 289–301.

Wang, S., Jiang, M., Duchesne, X. M., Laugeson, E. A., Kennedy, D. P., Adolphs, R., et al. (2015). Atypical Visual Saliency in Autism Spectrum Disorder Quantified through Model-Based Eye Tracking. *Neuron*, *88*, 1–13.

Wass, S. V., & Smith, T. J. (2014). Individual Differences in Infant Oculomotor Behavior During the Viewing of Complex Naturalistic

Scenes. *Infancy*, *19*(4), 352–384.

Wheeler, A., Anzures, G., Quinn, P. C., Pascalis, O., Omrin, D. S., & Lee, K. (2011). Caucasian Infants Scan Own- and Other-Race Faces Differently. *PLoS ONE*, *6*(4), e18621.

Winkler, S., & Subramanian, R. (2013). Overview of Eye Tracking Datasets. In *Qomex* (pp. 212–217). Klagenfurt, Austria.

Yarbus, A. L. (1965). *Eye Movements and Vision*. New York, USA: Plenum Press.

Yovel, G., Tambini, A., & Brandman, T. (2008). The asymmetry of the fusiform face area is a stable individual characteristic that underlies the left-visual-field superiority for faces. *Neuropsychologia*, *46*(13), 3061–3068.

Zhong, M., Zhao, X., Zou, X.-c., Wang, J. Z., & Wang, W. (2014). Markov chain based computational visual attention model that learns from eye tracking data. *Pattern Recognition Letters*, *49*, 1–10.