

**Human spontaneous gaze patterns in viewing of faces of
different species**

Kun Guo, David Tunnicliffe, Hettie Roebuck

Research Centre for Comparative Cognition, School of Psychology,

University of Lincoln, Lincoln LN6 7TS, UK

Corresponding Author:

Dr. Kun Guo

School of Psychology, University of Lincoln, Lincoln, LN6 7TS, UK

Email address: kguo@lincoln.ac.uk

Tel: +44-1522-886294

Keywords: Gaze pattern, Fixation, Faces, Human, Species

Abstract

Human studies have reported clear differences in perceptual and neural processing of faces of different species, implying the contribution of visual experience to face perception. Can these differences be manifested in our eye scanning patterns while extracting salient facial information? Here we systematically compared non-pet owners' gaze patterns while exploring human, monkey, dog and cat faces in a passive viewing task. Our analysis revealed that the faces of different species induced similar patterns of fixation distribution between left and right hemi-face, and among key local facial features with the eyes attracting the highest proportion of fixations and viewing times, followed by the nose and then the mouth. Only the proportion of fixation directed at the mouth region was species-dependent and could be differentiated at the earliest stage of face viewing. It seems that our spontaneous eye scanning patterns associated with face exploration were mainly constrained by general facial configurations; the species affiliation of the inspected faces had limited impact on gaze allocation, at least under free viewing conditions.

Introduction

Human faces are probably the most important visual stimuli in our social environment, and the processing of faces seems to involve a face-specific cognitive and neural mechanism (McKone et al 2006; see also Tarr and Cheng 2003). For instance, behavioural and neuropsychological studies have observed detrimental recognition performance for inverted faces rather than non-face objects (face inversion effect; e.g. Valentine 1988), and selective impairments of face and object recognition in neurological

patients (prosopagnosia and visual agnosia) (Farah 1996; Moscovitch et al 1997). Brain imaging studies further suggested a distributed network of brain structures, including the fusiform gyrus, associated with face processing (Gobbini and Haxby 2007).

This ‘special’ processing for faces seems to be species and race-sensitive, and is likely associated with our extensive experience of identifying conspecific faces. Developmental studies revealed that 6-month-old infants perform equally well at discriminating individual human or monkey faces; 9-month-olds, like adults, show better performance for recognizing frequently-experienced human faces (Pascalis et al 2002), suggesting a perceptual narrowing process in the development of our highly efficient face perception. The holistic or configural processing for faces (i.e. perceiving relations among facial features and integrating all features into an individual representation of the face as a whole. The configural processing is often assessed by face inversion effect which is defined as a larger decrease in recognition performance for faces than for other mono-oriented objects when they are presented upside-down) is more evident with faces of own species or even own race than with faces of other species or other races (McKone et al 2006; Rhodes et al 2006; Tanaka et al 2004), implying differences in the perceptual processing of faces from different species and races. Human brain imaging studies further revealed brain waveform (i.e. face-specific N170 event-related potential component) differences in latency, amplitude and distribution in response to faces of humans and animals (i.e. apes, dogs, cats and birds) (Bentin et al 1996; Carmel and Bentin 2002; de Haan et al 2002; Rousselet et al 2004), and cortical face-selective region activation differences in amplitude in response to faces of different races (Golby et al 2001),

suggesting that our visual experience with different face types have profound effects on the neural processing of facial information.

Could reported differences in the perceptual and neural processing of different face types be manifested in our gaze patterns associated with face viewing? In other words, do we employ a single/general oculomotor strategy to extract salient/relevant facial information from faces sharing similar spatial configurations (i.e. two eyes above a nose and a mouth)? Given the pattern of our eye movements can be modulated by cognitive demands and characteristics of the observed scenes (Guo et al 2006; Henderson 2003), it is reasonable to assume that our gaze patterns would be sensitive to the species-specific facial information. Here we systematically compared human participants' gaze patterns while free-viewing faces of humans, monkeys, dogs and cats, and observed similar viewing patterns to different face types, suggesting the dominant role of general facial configuration in shaping of face-related eye scanning patterns. The species of the inspected faces, on the other hand, had limited impact on gaze allocation, at least in the adopted free-viewing tasks.

While free-viewing conspecific faces, human adults often demonstrate a natural gaze bias towards the left visual field, that is, the right side of the viewee's face is often inspected first and/or for longer periods (Butler et al 2005; Mertens et al 1993; Philips and David 1997). This left gaze bias in face exploration in healthy observers is related to neither handedness nor eye dominance (e.g. Leonards and Scott-Samuel 2005). Although human visuospatial attention bias is to the left visual field and in some cultures, a long practised left-to-right directional scanning bias (most notably, reading) may contribute to this gaze asymmetry (e.g. Nicholls and Roberts 2002; Vaid and Singh 1989), it is often

argued that a right hemisphere advantage in face processing (receiving visual input from left visual field) is the likely cause of this asymmetry (Burt and Perrett 1997; Butler et al. 2005). A recent recording of human saccadic eye movements further revealed that the initial gaze bias is the most evident while exploring upright faces, and is less or not evident while exploring inverted faces and symmetric non-face object or landscape images (Leonards and Scott-Samuel 2005), suggesting this gaze asymmetry is part of gaze patterns associated with face exploration.

Interestingly, this face-related left gaze bias is not restricted to humans, but also occurs in non-human species such as rhesus monkeys (*Macaca mulatta*) and domestic dogs (*Canis familiaris*). In some animals such gaze asymmetry is even species-sensitive. For instance, domestic pet dogs only demonstrate a left gaze bias towards human faces, but not towards monkey or dog faces, implying a broader adaptive value of the left gaze bias in social species (Guo et al 2009). Given these findings, it would be interesting to examine whether we have similar pattern of gaze asymmetry while viewing faces of different species. If the left gaze bias is “an automatic, internally driven initiation of the saccadic exploration of faces” in humans (Leonards and Scott-Samuel 2005, p2679), then we are likely to demonstrate the same pattern of gaze asymmetry while exploring faces sharing similar spatial configurations. If, on the other hand, the left gaze bias is associated with processing species-sensitive facial information (Guo et al 2009), then we could show different degree of gaze asymmetry while inspecting faces of different species. In this study, our participants demonstrated a constant initial left gaze bias while free-viewing human, monkey, dog and cat faces, suggesting a general oculomotor

strategy to sample facial information across different species, at least for those sharing similar facial configurations.

Materials and methods

28 undergraduate psychology students (8 male, 20 female), age ranging from 19 to 44 years old with the mean of 22 ± 5.4 (Mean \pm SD), volunteered to participate in the study in return for course credit. All participants had normal visual acuity and were chosen from non-pet owners for monkeys, dogs and cats (to avoid potential influence of visual expertise effect; e.g. Tarr and Cheng 2003). Informed consent was obtained from each participant, and all procedures complied with the British Psychological Society “Code of Ethics and Conduct”, and with the World Medical Association Helsinki Declaration as revised in October 2008.

Digitized grey scale face images were presented through a ViSaGe graphics system (Cambridge Research Systems) and displayed on a high frequency non-interlaced gamma-corrected color monitor (30.0 cd/m^2 background luminance, 100 Hz frame rate, Mitsubishi Diamond Pro 2070SB) with the resolution of 1024×768 pixels. At a viewing distance of 57 cm the monitor subtended a visual angle of $40 \times 30^\circ$.

Four different categories of unfamiliar face images with closed mouth and neutral facial expressions in full frontal view were used as stimuli (see examples in Fig.1): 10 human faces, 10 monkey faces, 10 dog faces and 10 cat faces. All images shared similar spatial facial configurations, were gamma-corrected and displayed once in a random order at the centre of the screen with a resolution of 600×600 pixels ($22 \times 22^\circ$).

During the experiments the participants sat in a chair with their head restrained by a chin rest, and viewed the display binocularly. To calibrate eye movement signals, a small red fixation point (FP, 0.3° diameter, 15 cd/m² luminance) was displayed randomly at one of 9 positions (3 × 3 matrix) across the monitor. The distance between adjacent FP positions was 10°. The participant was instructed to follow the FP and maintain fixation for 1 sec. After the calibration procedure, the trial was started with a FP displayed on the centre of the monitor. If the participant maintained fixation for 1sec, the FP disappeared and an image was presented for 3 sec. During the presentation, the participant passively viewed the images with the instruction of “viewing the faces as you normally do”. No reinforcement was given during this procedure. It was considered that in the absence of instrumental responding, our participants’ viewing behaviour should be as natural as possible. The inter-trial interval was set to 2 sec.

Horizontal and vertical eye positions were measured using a Video Eyetracker Toolbox with 50 Hz sampling frequency and up to 0.25° accuracy (Cambridge Research Systems). The software developed in Matlab computed horizontal and vertical eye displacement signals as a function of time to determine eye velocity and position. Fixation locations were then extracted from the raw eye tracking data using velocity (less than 0.2° eye displacement at a velocity of less than 20°/s) and duration (greater than 50 ms) criteria (Guo et al 2006).

While determining fixation allocation within key internal facial features (i.e. eyes, nose and mouth), we adopted the criteria from Barton et al (2006) to consistently define boundaries between local facial features for different faces (for an example of defining human facial regions, see <http://www.perceptionweb.com/perception/misc/p5547/f3.jpg>).

Specifically, the ‘eye’ region included the eyes, eyelids and eye brows; the ‘nose’ or ‘mouth’ region consist of main body of the nose (glabella, nasion, tip-defining points, alar-sidewall and supra-alar crease) or mouth (only lips were visible for closed mouths) and immediate surrounding area (up to 0.5° visual angle). The division line between the ‘mouth’ and ‘nose’ regions was the midline between upper lip and the bottom of the nose. Each fixation was then characterised by its location among feature regions and its time of onset relative to the start of the trial, and the number of fixations directed at each facial feature was normalized to the total number of fixations sampled in that trial. As the same facial feature across faces of different species often vary in size (i.e. dogs usually have larger noses than humans), the proportion of the area of a particular facial feature relative to the whole image was subtracted from the proportion of fixations directed at that facial feature in a given trial. Any difference in fixation distribution from zero means that this particular facial feature attracted more or less fixations than predicted by a uniform looking strategy (Dahl et al 2009).

Results

Faces of different species presented in the free viewing tasks attracted similar amount of attention from our participants. One way repeated measures analysis of variance (ANOVA) showed non-significant differences in the number of fixations per image across human (7.91 ± 0.3 , Mean \pm SEM), monkey (7.52 ± 0.33), dog (8.03 ± 0.32) and cat faces (7.65 ± 0.26) ($F(3,108) = 0.59$, $p = 0.62$).

--- Figure 1 about here ---

Analysis of fixation allocation revealed that immediately following the face presentation, the first saccade was directed at the eye region in $87\% \pm 2$ of the trials

(averaged across participants); and during the face exploration, the vast majority of fixations ($91\% \pm 1$ of overall fixations) and viewing time ($91\% \pm 0.4$ of total face viewing time within a trial) were allocated at key internal facial features, such as eyes, nose and mouth. We then examined whether faces of different species attracted similar distribution of fixation and viewing time across these local facial features (Fig. 1). As the experimental design comprised four levels of face types (human, monkey, dog and cat faces), three levels of local features (eyes, nose and mouth) and two dependent variables (distribution of fixation and viewing time within faces), two two-way repeated measures ANOVA was carried out after averaging the proportion of fixations directed at each local feature (the number of fixations within each facial feature as percentage of total number of fixations within whole face image subtracting the proportion of the area of each facial feature relative to the whole image) for each face type and each participant, and after averaging the proportion of viewing time directed at each local feature (cumulative viewing time within each facial feature as percentage of total face viewing time subtracting the proportion of the area of each facial feature relative to the whole face) for each face type and each participant.

Our analysis showed that qualitatively, faces of different species attracted similar patterns of fixation distribution ($F(3,324)=0.61$, $p=0.61$) and viewing time distribution ($F(3,324)=0.43$, $p=0.73$) across local facial features with the eyes attracting the highest proportion of fixations (51–59%) and viewing time (54–59%), followed by the nose (fixation 11–14%, viewing time 10–13%) and then the mouth (fixation 1–7%, viewing time 1–8%) (fixation distribution across eyes, nose and mouth: $F(2,324)=742$, $p<0.001$; viewing time distribution across eyes, nose and mouth: $F(2,324)=781$, $p<0.001$; Fig. 1).

However, the significant interaction effect between local features and face types (fixation $F(6,324)=3.27$, $p=0.004$; viewing time $F(6,324)=2.89$, $p=0.01$) suggested that the quantitative distribution of fixations and viewing time within some facial features was species-dependent. Specifically, human observers paid the same amount of attention to the eyes or nose region within faces of different species (Bonferroni correction for multiple comparisons, $p>0.22$). They also directed indistinguishable proportions of fixations and viewing times at the mouth region in human and monkey faces ($p>0.18$), but significantly less fixations and viewing times at the mouth region in dog and cat faces in comparison with the mouth region in both human and monkey faces ($p<0.001$).

--- Figure 2 about here ---

We further examined when this differential gaze allocation to the mouth region of faces of different species happened during face exploration. Given that local image regions scoring high on saliency or relevancy measures are likely to be inspected earlier in picture viewing (e.g. Parkhurst and Niebur 2003), analyzing sequential fixation placement in faces could provide valuable relevancy information about individual facial features in the face processing. In this study the probabilities of sequential fixation placement in the mouth for each of the first five fixations sampled in face viewing were compared across different faces types (Fig. 2). We chose to analyze the first five sequential fixation placements because our participants made at least 5 fixations in the majority (>92%) of the trials, and the initial time window of face viewing is critical in face processing (Dahl et al 2009). For example, a recent study by Hsiao and Cottrell (2008) revealed that the first two fixations in face viewing were sufficient to achieve optimal face recognition performance.

The 4 (face type) \times 5 (fixation sequence) repeated measures ANOVA revealed a significant main effect of face type ($F(3,540)=44.89$, $p<0.001$; Fig. 2A) and fixation sequence ($F(4,540)=16.39$, $p<0.001$), and significant interaction effect between face type and fixation sequence ($F(12,540)=2.37$, $p=0.006$). Specifically, for the first fixation, the mouth in human faces was more likely to be inspected than the mouth in dog and cat faces. For the next four fixations, the mouth in both human and monkey faces had higher chance to be the saccade target than that in dog and cat faces (Bonferroni correction for multiple comparisons, $p<0.01$). Given that the variance in mouth size across different species could bias our conclusion (i.e. large mouth could have higher chance to be saccade destination), we performed the same analysis after normalising the probability of sequential fixation placement in the mouth region according to its size proportion relative to the whole image. For non-human faces, the probability of the mouth to be fixated was divided by the ratio of its size proportion relative to the human mouth (the size of the human mouth was treated as 100%, Fig. 2B). Two way ANOVA and associated Bonferroni post-hoc tests reached similar conclusions as the prior-normalisation analysis. It seemed that gaze allocation to the mouth region was species-dependent and could differentiate at the earliest stage of face viewing.

--- Figure 3 about here ---

To examine whether the faces of different species could induce different pattern of gaze asymmetry in humans, we also compared the probability of sequential fixations directed at the left and right hemi-face of each viewed face types (all fixations to left and right hemi-face were included, including those outside the eye, nose and mouth regions). Compared with the right hemi-face, the left hemi-face (from viewer's perspective) on

average attracted a higher probability of the first gaze direction after image presentation (72-78% across face types, Fig. 3A). 2 (left and right hemi-face) \times 4 (face type) repeated measures ANOVA revealed a significant main effect of face side ($F(1,216)=175.46$, $p<0.001$), but non-significant main effect of face type ($F(3,216)=0$, $p=1$) and interaction between face side and face type ($F(3,216)=0.78$, $p=0.51$). For the next four fixations, the left and right hemi-face had the same probabilities to be fixated regardless of species (Fig. 3B). 4 (face type) \times 5 (fixation sequence) repeated measures ANOVA of the probability of sequential fixation directed at the left hemi-face revealed a significant main effect of fixation sequence ($F(4,540)=30.47$, $p<0.001$), but non-significant main effect of face type ($F(3,540)=1.35$, $p=0.25$) and interaction between face type and fixation sequence ($F(12,540)=0.59$, $p=0.85$). Clearly, the left gaze bias we observed here was the most evident for the initial fixation, but was not species-sensitive.

So far our analysis has revealed that the faces of different species induced similar patterns of gaze distribution. Only the proportion of fixation and viewing time directed at the mouth region was species-dependent and could be differentiated at the earliest stage of face viewing. It could be argued that the differences in gaze allocation to the mouth across face categories were driven by low-level image salience (i.e. local image contrast, intensity and structure) rather than anything category-specific. To examine this possibility, we calculated the top eight salient regions within each face image using the most widely used saliency model of Itti and Koch (2000), with the authors' original parameters and implementation (obtained from <http://ilab.usc.edu>). The model compares local image intensity, colour and orientation, combines them into a single saliency map with a winner-take-all network and inhibition-of-return, and then produces a sequence of

predicted fixations that scan the scene in order of decreasing saliency. We chose to calculate the first eight salient regions within the image because our participants on average made between 7 and 8 fixations per images in face viewing.

Out of 10 images per face category, no human mouth could be classified within the top eight salient image regions. The mouth in two monkey faces and one dog face was ranked as the third salient region within the image, and only one cat mouth was ranked as the eighth salient image region. If our gaze allocation to the mouth was purely driven by the local image salience, then the mouth in monkey faces should attract more fixations than the mouth in dog and cat faces. Human mouth, on the other hand, should not attract any fixations at all. As can be seen in Fig. 1A, our participants' gaze allocations to the mouth region in viewing faces of different species were completely different from those predicted by the saliency map. It seems that the higher proportion of fixations and viewing times directed at human and monkey mouths could not be fully accounted for by low-level local image salience.

Discussion

Viewing of conspecific faces in humans is associated with a stereotypical pattern of eye movements. Among various internal and external local facial features, eyes, nose and mouth attract the majority of fixations with the eye region being inspected first and most frequently (Althoff and Cohen 1999; Barton et al 2006; Guo et al 2003; Heisz and Shore 2008; Stacey et al 2005). In this study we extended this finding to faces of different species, including less frequently encountered face types (i.e. monkey faces). Our overall viewing patterns to human, monkey, dog and cat faces were almost identical. Regardless

of species, the eyes were often inspected first and attracted the highest proportion of fixations and viewing times, followed by the nose and then the mouth region. Furthermore, a consistent left gaze bias was associated with the initial stage of face exploration irrespective of face types. It seems that under free-viewing conditions we tend to use a general oculomotor strategy to sample facial information across different species, at least for those sharing similar facial configurations (i.e. vertical bilateral symmetry of the spatial arrangement of two eyes above a nose and a mouth). In other words, human-face-like visual images would trigger stereotypical eye scanning pattern automatically, regardless of observer's perceptual expertise/experience.

Recent behavioural and neurophysiological studies suggest a species-sensitive face processing. The differences in perceptual and cognitive processing of different face types, however, are more quantitative rather than qualitative. For instance, although human observers perform better at differentiating human faces which involves a more holistic processing strategy (i.e. more evident face inversion effect for human faces than monkey faces; McKone et al 2006; Mondloch et al 2006), human and monkey faces share very similar facial configurations and can be categorized into the same perceptual group in category identification tasks (Campbell et al 1997). Furthermore, the face-specific N170 ERP component elicited in humans by human faces is as large and distinctive as that elicited by monkey faces, only peaks up to 10 ms earlier (Carmel and Bentin 2002; Rousselet et al. 2004). It seems that although our face perception is species-sensitive, the faces from those species sharing similar facial configurations have limited influence on the face processing.

Our quantitative comparison of gaze allocation to individual facial features also demonstrated a limited influence of face types on fixation distribution. After adjusting for the variance in size of local facial features across species, our analysis revealed that the proportion of fixations and viewing times directed at the eyes or nose region were indistinguishable across human, monkey, dog and cat faces. The mouth region in human and monkey faces, on the other hand, attracted significantly more fixations and longer viewing times than that in dog and cat faces. Analysis of sequential fixation placement in the mouth region further revealed that such difference in gaze allocation started to differentiate at the earliest stage of face exploration, namely from the first fixation. As different facial features could provide different types and amounts of facial information (i.e. the eyes contain critical information about face identity and social attention, the mouth is crucial for fast detection and recognition of some facial expressions) (Heisz and Shore 2008; Schyns et al 2007), the differential gaze allocation to the mouth could reflect a different viewing strategy/sensitivity to sample relevant facial information from different species. Given the relevance and importance of the human mouth in transforming a range of diagnostic expression cues in our social communication, and the high similarity in spatial configuration between human and monkey faces, it is quite possible that during free exploration we involuntarily direct a substantial amount of attention to human and monkey mouths to evaluate subtle expression and emotion cues. Our participants (non-pet owners) did not engage similar gaze distribution in the viewing of dog and cat faces as they may not have interest and/or perceptual experience in processing subtle emotion cues from dog and cat mouths. It would be interesting to

address this possibility by comparing gaze patterns in the viewing of dog/cat faces between pet owners and non-pet owners.

It could be argued that the difference in the proportion of fixations directed at the mouth region of different face types is due to the differences in local image structural properties (i.e. size, local contrast and local image structure). Although it is difficult to control these variables in realistic face photos, our previous studies have revealed that the physical properties of local facial features cannot account for normal fixation distribution within the faces. Taking the eye region as an example, for faces used in this study the eye region on average only occupied 7-9% of image size, but attracted disproportionately 58-68% of total fixations. Changing its location or surrounding context but keeping intact local structure and contrast (i.e. by scrambling faces) would significantly reduce the number of attracted fixations (Guo et al 2003, 2007). Furthermore, local facial regions with high image salience (based on the calculation of local image physical properties) are not necessarily correlated with the gaze distribution in face viewing. For example, with relatively high local contrasts and complex local structures (higher spatial frequency and frequent variances in local orientation/curvature), regions of human hairline are often regarded as the most salient facial regions by the classical saliency models such as the one proposed by Itti and Koch (2000), but they received few, if any fixations from our participants during face exploration. Hence the gaze distribution within a face is more likely dependent upon the amount of available facial information contained within each facial feature, rather than constrained by their simple physical properties.

Taken together, it seems that our spontaneous viewing pattern in face exploration is largely constrained by the facial configurations. But during the course of exploration,

the prior knowledge/experience about certain face types could influence the detailed distribution of fixations directed at the mouth region. The study of gaze pattern, therefore, may help to reveal the effect of such prior knowledge or other semantic factors on perceptual processing of faces.

It remains to be seen to what extent our findings can be generalized to different task contexts. By the presentation of static face images or dynamic video recordings, previous studies have suggested a modulatory role of task-based top-down guidance in determining fixation allocation in face exploration (e.g. Buchan et al 2000; Malcolm et al 2008). For example, while judging which of the two simultaneously presented human faces was similar to a prior presented face in visual appearance, identity or expression, participants made a significant shift from more scanning of upper- than lower-face in identity judgements, to more scanning of lower- than upper-face in expression judgements. No such shift was observed in judging appearance similarity for faces differed on identity or expression, suggesting a top-down task effect on gaze behaviour (Malcolm et al 2008). It will be interesting to examine whether we adopt similar task-dependent gaze patterns while processing specific facial information from non-human faces.

Reference

- Althoff RR, Cohen NJ, 1999 “Eye-movement-based memory effect: A reprocessing effect in face perception” *Journal of Experimental Psychology: Learning, Memory, and Cognition* **25** 997-1010
- Barton JJS, Radcliffe N, Cherkasova MV, Edelman J, Intriligator JM, 2006 “Information processing during face recognition: the effects of familiarity, inversion, and morphing on scanning fixations” *Perception* **35** 1089-1105
- Bentin S, Allison T, Puce A, Perez A, McCarthy G, 1996 “Electrophysiological studies of face perception in humans” *Journal of Cognitive Neuroscience* **8** 551–565

- Buchan JN, Pare M, Munhall KG, 2000 “Spatial statistics of gaze fixations during dynamic face processing” *Social Neuroscience* **2** 1-13
- Burt DM, Perrett DI, 1997 “Perceptual asymmetries in judgements of facial attractiveness, age, gender, speech and expression” *Neuropsychologia* **35** 685-693
- Butler S, Gilchrist ID, Burt DM, Perrett DI, Jones E, Harvey M, 2005 “Are the perceptual biases found in chimeric face processing reflected in eye-movement patterns?” *Neuropsychologia* **43** 52-59
- Campbell R, Pascalis O, Coleman M, Wallace SB, Benson PJ, 1997 “Are faces of different species perceived categorically by human observers?” *Proceedings of the Royal Society of London, Series B* **264** 1429-1434
- Carmel D, Bentin S, 2002 “Domain specificity versus expertise: factors influencing distinct processing of faces” *Cognition* **83** 1-29
- Dahl CD, Wallraven C, Bühlhoff HH, Logothetis NK, 2009 “Humans and macaques employ similar face-processing strategies” *Current Biology* **19** 509-513
- de Haan M, Pascalis O, Johnson M, 2002 “Specialization of neural mechanisms underlying face recognition in human infants” *Journal of Cognitive Neuroscience* **14** 199-209
- Farah MJ, 1996 “Is face recognition ‘special’? Evidence from neuropsychology” *Behaviour Brain Research* **76** 181-189
- Gobbini MI, Haxby JV, 2007 “Neural systems for recognition of familiar faces” *Neuropsychologia* **45** 32-41
- Golby AJ, Gabrieli JDE, Chiao JY, Eberhardt JL, 2001 “Differential responses in the fusiform region to same-race and other-race faces” *Nature Neuroscience* **4** 845-850
- Guo K, Robertson RG, Mahmoodi S, Tadmor Y, Young MP, 2003 “How do monkeys view faces? – A study of eye movements” *Experimental Brain Research* **150** 363-374
- Guo K, Mahmoodi S, Robertson RG, Young MP, 2006 “Longer fixation duration while viewing face images” *Experimental Brain Research* **171** 91-98
- Guo K, 2007 “Initial fixation placement in face images is driven by top-down guidance” *Experimental Brain Research* **181** 673-677
- 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
- Heisz JJ, Shore DI, 2008 “More efficient scanning for familiar faces” *Journal of Vision* **8** 1-10
- Henderson JM, 2003 “Human gaze control during real-world scene perception” *Trends in Cognitive Sciences* **7** 498-504
- Hsiao JH, Cottrell GW, 2008 “Two fixations suffice in face recognition” *Psychological Science* **9** 998-1006
- Itti L, Koch C, 2000 “A saliency-based search mechanism for overt and covert shifts of visual attention” *Vision Research* **40** 1489-1506
- Leonards U, Scott-Samuel NE, 2005 “Idiosyncratic initiation of saccadic face exploration in humans” *Vision Research* **45** 2677-2684
- Malcolm GL, Lanyon LJ, Fugard AJB, Barton JJS, 2008 “Scan patterns during the processing of facial expression versus identity: An exploration of task-driven and stimulus-driven effects” *Journal of Vision* **8(8):2** 1-9
- McKone E, Kanwisher N, Duchaine BC, 2006 “Can generic expertise explain special processing for faces?” *Trends in Cognitive Sciences* **11** 8-15

- Mertens I, Siegmund H, Grusser OJ, 1993 "Gaze motor asymmetries in the perception of faces during a memory task" *Neuropsychologia* **31** 989-998
- Mondloch CJ, Maurer D, Ahola S, 2006 "Becoming a face expert" *Psychological Science* **17** 930-934
- Moscovitch M, Winocur G, Behrmann M, 1997 "What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition" *Journal of Cognitive Neuroscience* **9** 555-604
- Nicholls MER, Roberts GR, 2002 "Can free-viewing perceptual asymmetries be explained by scanning, pre-motor or attentional biases?" *Cortex* **38** 113-136
- Parkhurst DJ, Niebur E, 2003 "Scene content selected by active vision" *Spatial Vision* **16** 125-154
- Pascalis O, de Haan M, Nelson CA, 2002 "Is face processing species-specific during the first year of life?" *Science* **296** 1321-1323
- Philips ML, David AS, 1997 "Viewing strategies for simple and chimeric faces: An investigation of perceptual bias in normal and schizophrenic patients using visual scan paths" *Brain and Cognition* **32** 225-238
- Rhodes G, Hayward WG, Winkler C, 2006 "Expert face coding: Configural and component coding of own-race and other-race faces" *Psychonomic Bulletin & Review* **13** 499-505
- Rousselet GA, Mace MJM, Fabre-Thorpe M, 2004 "Animal and human faces in natural scenes: how specific to human faces is the N170 ERP component?" *Journal of Vision* **4** 13-21
- Schyns PG, Petro LS, Smith ML, 2007 "Dynamics of visual information integration in the brain for categorizing facial expressions" *Current Biology* **17** 1580-1585
- Stacey PC, Walker S, Underwood JDM, 2005 "Face processing and familiarity: evidence from eye-movement data" *British Journal of Psychology* **94** 407-422
- Tanaka J, Kiefer M, Bukach CM, 2004 "A holistic account of the own-race effect in face recognition: evidence from a cross-cultural study" *Cognition* **93** 1-9
- Tarr MJ, Cheng YD, 2003 "Learning to see faces and objects" *Trends in Cognitive Sciences* **7** 23-30
- Vaid J, Singh M, 1989 "Asymmetries in the perception of facial affect: Is there an influence of reading habits?" *Neuropsychologia* **27** 1277-1287
- Valentine T, 1988 "Upside-down faces: a review of the effects of inversion upon face recognition" *British Journal of Psychology* **79** 471-491

Figure 1

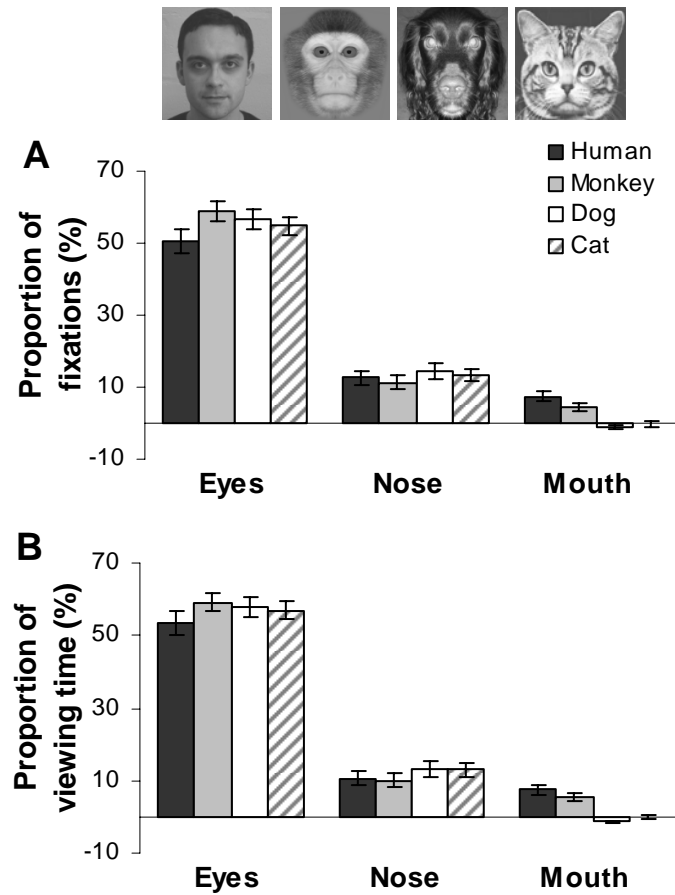


Figure 1, (A) Number of fixations directed at eyes, nose and mouth regions as percentage of total number of fixations within whole face image of different species (human, monkey, dog and cat faces). (B) Cumulative viewing time directed at eyes, nose and mouth regions as percentage of total face viewing time. The proportion of the area of each facial region relative to the whole image was subtracted from the proportion of the fixations directed at each corresponding facial part. Any difference in fixation distribution from zero means that this particular facial region was inspected more or less than predicted by a uniform looking strategy. Errors bars indicate standard error of mean.

Figure 2

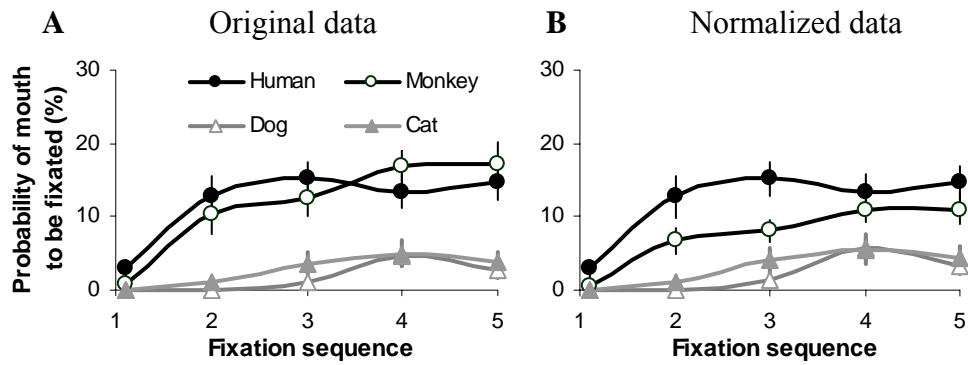


Figure 2, (A) The probability of the mouth region as the destination of first five fixations while free-viewing faces of different species. (B) Normalised probability of sequential fixation placement in the mouth region according to its size proportion relative to the whole image. For non-human faces, the probability of the mouth to be fixated was divided by the ratio of its size proportion relative to human mouth (the size of human mouth was treated as 100%). Errors bars indicate standard error of mean.

Figure 3

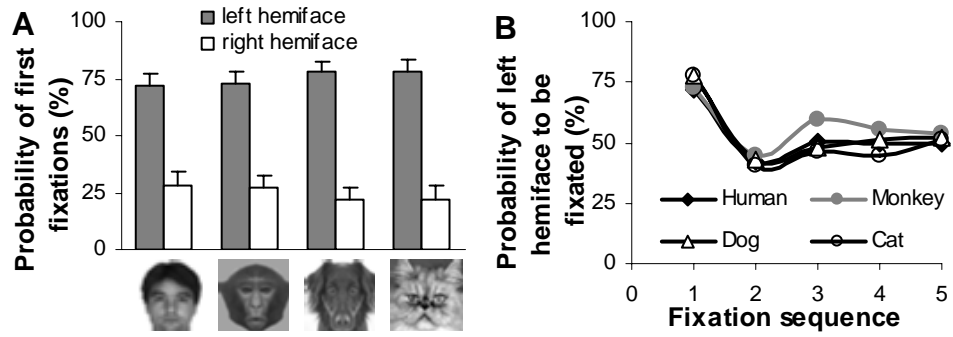


Figure 3, (A) The probability of initial fixation directed at left and right side of presented faces. Errors bars indicate standard error of mean. (B) The probability of sequential fixation directed at the left side of presented faces of different species.