Vision Research 157 (2019) 222–229

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

The adult face-diet: A naturalistic observation study

Ipek Oruc[a,b,*], Fakhri Shafai[a,b], Shyam Murthy[a], Paula Lages[a], Thais Ton[a]

[a] Department of Ophthalmology and Visual Sciences, University of British Columbia, Canada
[b] Graduate Program in Neuroscience, University of British Columbia, Canada

A R T I C L E I N F O

No of reviewers – 1

Keywords:
Face-diet
Exposure
Experience
Other-race effect
Face inversion effect
View-point invariance
Three-quarter view advantage
Face size
Naturalistic observation
Environment
Evolution
First-person perspective

A B S T R A C T

Experience plays a fundamental role in the development of visual function. Exposure to different types of faces is an important factor believed to shape face perception ability. Contents of daily exposure to faces, i.e., the face-diet, of infants have been documented in previous studies. While face perception involves a protracted development and continues to be malleable well into adulthood, an empirical study of the adult face-diet has been lacking. We collected first-person perspective footage from 30 adults during the course of their daily activities. We found that adults' exposure to faces is longer and more diverse compared to that of infants. Frequency of exposure were highest for familiar (75%), own-race (81%), and three-quarter pose (44%) faces. Faces in the adult face-diet were relatively large (median 6°) suggesting fairly close viewing distances. Face sizes were significantly larger for familiar (median 7.1°) compared to unfamiliar (median 4.9°) faces, reflecting the closer viewing distances that characterize social interaction. These results are consistent with the view that face recognition processes are tuned to the ecologically relevant values of face attributes that are encountered most frequently in the real-life context to optimize face perception abilities.

1. Introduction

Face perception plays a key role in an individual's ability to form and maintain social interactions. Carrying out a face-to-face conversation is difficult without being able to interpret a friend's facial expressions; and failing to recognize a new colleague you met previously at a social event prevents a new connection to be made. The role of face recognition is undeniable in navigating one's social environment. Recognizing faces differs from most other visual object recognition tasks since it involves individuating exemplars that share the same configuration (two eyes above the nose above the mouth) based on subtle differences e.g., between features and relative distances between features. Human observers are considered to be experts in face perception as they maintain sensitivity to such subtle differences that distinguish different identities while remaining robust across significant changes among images of the same identity, e.g., due to changes in viewing conditions such as lighting and viewpoint. Despite the apparent complexity of this task, large numbers of faces encountered in a lifetime are remembered and recognized by human observers with seemingly little difficulty.

How human observers develop face expertise continues to be a topic of controversy. Some evidence is suggestive of innate and genetic contributions (see McKone, Crookes, Jeffery, & Dilks, 2012; McKone, Kanwisher, & Duchaine, 2007, for reviews). For example, Farah and colleagues describe the case of a classic presentation of acquired prosopagnosia resulting from brain damage suffered only one day after birth (Farah, Rabinowitz, Quinn, & Liu, 2000) suggesting that the neuroanatomical structures devoted to face processing may be genetically predetermined independent of visual experience. Furthermore, multiple studies show that newborns orient toward faces and face-like forms prior to any significant opportunities for visual experience (Goren, Sarty, & Wu, 1975; Johnson, Dziurawiec, Ellis, & Morton, 1991), suggesting a built-in predisposition for responding to faces. In addition, twin studies of face recognition show that this specialized ability is heritable providing evidence for a genetic basis for face processing (Wilmer et al., 2010; Zhu et al., 2010).

On the other hand, it is clear that innate face knowledge cannot fully explain human face recognition abilities. Conclusive evidence for this comes from the so-called other-race effect—a robust finding of diminished recognition and memory for other-race faces, compared to those of own-race (Hayward, Rhodes, & Schwaninger, 2008; Meissner & Brigham, 2001; O'Toole, Deffenbacher, Valentin, & Abdi, 1994; Rhodes, Hayward, & Winkler, 2006; Rostamirad, Barton, & Oruc, 2009; Walker & Tanaka, 2003). The other-race effect is not a consequence of genetic determinants of ethnicity. For example, Korean children adopted by Caucasian European families between the ages of three to nine acquired

* Corresponding author at: University of British Columbia, Department of Ophthalmology and Visual Sciences, 818 West 10th Avenue, Vancouver B.C. V5Z 1M9, Canada.
E-mail address: ipor@mail.ubc.ca (I. Oruc).

https://doi.org/10.1016/j.visres.2018.01.001
Received 29 August 2017; Received in revised form 3 January 2018; Accepted 12 January 2018
Available online 01 February 2018

the reverse other-race effect, i.e., difficulties with recognizing East Asian faces (Sangrigoli, Pallier, Argenti, Ventureyra, & de Schonen, 2005). The fact that face expertise is fundamentally shaped and limited by chance events such as place of birth firmly establishes the key role played by visual experience.

Development of various visual processes involve sensitive periods during which appropriate visual input is required for normal development (see Lewis & Maurer, 2005, for a review). Similarly, early visual exposure may play a key role in the development of face expertise. Several studies show permanent deleterious effects of early visual deprivation from dense congenital cataracts on expert face processing (de Heering & Maurer, 2014; Geldart, Mondloch, Maurer, De Schonen, & Brent, 2002; Le Grand, Mondloch, Maurer, & Brent, 2001, 2003, 2004). Importantly, brief post-natal visual deprivation periods appear to be sufficient to engender these permanent deficits despite many years of visual experience following cataract reversal.

Finally, several lines of research point to continued impact of visual exposure to faces well beyond an early critical period. Evidence suggesting that early visual experience is not sufficient to support normal face processing is provided by studies on the case of MM who was blind between the ages of 3 and 43. Subsequent to restoration of his vision, MM's face perception remained severely and permanently impaired in contrast to his relatively intact motion and simple form processing (Fine et al., 2003; Huber et al., 2015). Face recognition goes through a protracted developmental course extending to late childhood and adolescence (e.g., Carey, Diamond, & Woods, 1980; de Heering, Rossion, & Maurer, 2012). Further evidence of late maturation comes from a large scale study by Germine and colleagues (2011) who report peak recognition memory for faces occurs in early 30 s. A study on adults of Korean origin adopted as children by European families between the ages of 3 and 9 show that the other-race effect can be reversed in late childhood (Sangrigoli et al., 2005) indicating that although the other-race effect is in place within the first year of life (Kelly et al., 2007), significant plasticity of the face-specific visual processes is still observed well into late childhood. In addition, visual training in adulthood has been shown to significantly reduce the face inversion effect (Laguesse, Dormal, Biervoye, Kuefner, & Rossion, 2012) and the other-race effect (Tanaka, Heptonstall, & Hagen, 2013). Overall, these results suggest that face recognition processes remain malleable in adulthood and continue to be shaped based on the visual experience of the observers.

Face-diet, a term coined by Rhodes and colleagues, refers to the collection of faces encountered as part of one's day-to-day visual experience (Rhodes, Jeffery, Watson, Clifford, & Nakayama, 2003). Despite the central role face-diet plays in shaping face recognition processes, empirical studies of the face-diet have been lacking until very recently. Five recent studies examining the face-diet of infants aged 1–24 months have contributed to advancing our understanding of infants' real-life exposure to faces (Fausey, Jayaraman, & Smith, 2016; Jayaraman, Fausey, & Smith, 2015, 2017; Sugden, Mohamed-Ali, & Moulson, 2014; Sugden & Moulson, 2017). Yet, a comprehensive empirical assessment of the adult face-diet has yet to be undertaken. It is uncertain how much can be extrapolated from what is known regarding the infant face-diet since daily activities of adults, and consequently their visual experiences, may vary drastically from those of infants.

Adults typically engage in a larger variety of activities compared to infants related to various functions such as their occupation, procurement of basic necessities, entertainment, physical activities, self-care, and social activities. For example, an adult may travel by bus to their workplace, go to a restaurant for lunch, visit a coffee shop in the afternoon, run errands or go shopping in the market, and attend a fitness class at a community centre within the time frame of a typical day. These activities bring with them a greater variety of social encounters with not only family members (as is typical with infants) but also with those of varying familiarity such as friends, colleagues, acquaintances, shop keepers, fellow commuters, or strangers passing by on a street. Physical factors, such as mobility and the size of one's body in relation to those of others in the environment, can also impact characteristics of one's visual input. Adults are able to move about freely and have the ability to orient toward, approach, as well as distance themselves from others and view faces around a level line of sight.

In the present study, we examined the adult face-diet to determine exposure statistics with regards to attributes such as viewing distance, pose, gender, ethnicity and familiarity. For this purpose, we equipped 30 adult participants with an eyewear-embedded camera and recorded first-person perspective footage while the participants carried on with their regular daily activities. In light of these results we compare the adult face-diet to that of infants reported in the literature and reveal the ways in which these two differ. In addition, we consider whether and to what degree statistics of adult face exposure can account for robust behavioural findings in the face recognition literature. Finally, we discuss potential implications of our results that engender new hypotheses regarding key factors in the development of face expertise for future study.

2. Methods

2.1. Eyewear-embedded camera

The footage was acquired using a high-resolution 75° field-of-view eyewear-embedded camera, Pivothead Durango (http://www.pivothead.com/). The camera was set to time-lapse mode to capture still images at the rate of 1 shot/30 s at 3-megapixel resolution. We replaced the shades with clear lenses and connected the glasses to a pocket-sized external battery (Pivothead Power Pro Refuel 8000), which the participants carried near or on their person (see Fig. 1).

2.2. Participants

Thirty adults (14 females; mean age = 31.9 ± 8.4 years, range 20–54) participated in the study. Out of the 30 participants: 28 were Caucasian, one was African and one was Asian; 14 were female; 18 participants recorded footage on a workday. In terms of occupation, the participant group consisted of four researchers, five engineers, one youth leader, two teachers, eleven students, one unemployed, one accountant, one administrative coordinator, one managing director, one customer services specialist, one bookkeeper, and one lab manager. Average footage recorded per participants was just over 7 h and 26 min (range: 2 h 35 m–13 h 23 m) for a total duration of 209 h and 57.5 min (25,195 frames).

Fig. 1. Eyewear-embedded camera. Eyewear-embedded camera was used to capture still images at the rate of 2 shots/min throughout the course of one day. A pocket-sized external battery was connected to the camera, which was carried on or near the participant's person.

2.3. Procedure

Participants were given the glasses ahead of time and were asked to wear them during waking hours of one day. Participants were instructed to turn on the camera upon waking and go about their daily activities as usual. No additional action was required from the participants as the camera automatically captured a still image every 30 s throughout the day. At the conclusion of the recording day, participants were given the option of connecting the glasses to their personal computer to review their footage. This was done to provide the participants with the opportunity to remove any images of a private nature (e.g. bathroom visits). Participants also completed a post-participation questionnaire where they indicated their gender, age, ethnicity, occupation, whether the recording was done on a work-day (vs. non-work day), whether they removed the glasses for any period of time, and any additional comments they had. The protocol was approved by the review boards of the University of British Columbia and Vancouver Hospital, and informed consent was obtained in accordance with the principles in the Declaration of Helsinki.

2.4. Data analysis

2.4.1. Face detection

The footage was pre-processed for automated detection of faces via in-house Matlab scripts. This was followed by manual adjustment of automated detections in which bounding boxes were drawn around faces that were missed by the automated process, and false detections and redundant bounding boxes due to multiple automated detections of the same face were deleted such that there was one, and only one, bounding box around every face that was captured in the footage. Face images that appeared in media, (e.g., print, screen) were not coded.

2.4.2. Face annotation

Each individual participant's footage was manually annotated for the following attributes: gender (female, male), ethnicity (Caucasian, Asian, African, other), pose (frontal, three-quarters, profile, other), and familiarity (familiar, unfamiliar, unsure). Coders were instructed to categorize faces as: 'frontal pose', if both eyes were visible and at least some portion of both ears were visible; 'three-quarters', if both eyes and one ear was visible; and 'profile': if one eye and one ear was visible (in all cases, features were accepted as "visible" if they would have been visible if not for a reason other than head rotation, such as occlusion by hair). Coders annotated the photos in the same temporal order they were captured. The temporal continuation of the photos provided a context to the situations depicted in the still images and allowed the coders gain a better understanding of the events and activities captured in the footage. This was done to improve the coders' ability to interpret the images, and increase coders' accuracy and confidence on the judgements they make in the specified attributes. Inter-rater reliability was defined as the correlation coefficient between two independent coders' frequency estimates of a given attribute (e.g. female) across all participants.

All annotations, including familiarity, were based on the subjective judgments of the coders. Familiarity judgments were based on the perceived nature and setting of interaction between the participant and the individuals in the footage, as well as on the temporal continuation of the daily activities in the serially coded images, and consistent recurrence of specific individuals in activities that are indicative of familiarity. For example, a participant's footage may feature images of early morning activities at home, such as eating breakfast and interactions with family (based on this interpretation the coder would annotate family members as 'familiar'), followed by a commute to work by bus. If individuals in the bus did not appear to interact with the participant, then fellow commuters on the bus would all be coded as 'unfamiliar'. A face initially coded as 'unfamiliar' could be updated to 'familiar' if subsequent images in the footage suggested it. Coders were instructed to be conservative in their annotations and to use the 'unsure' option unless they were reasonably confident in their judgment. Examples of faces annotated as 'familiar' include family members who are in the house in the morning and evening; classmates who are interacting; instructors in the classroom; the person who shares a meal with the participant in a restaurant environment among others. Examples of faces annotated as 'unfamiliar' include passers-by on the street; fellow commuters on a bus; cashier in a store, and people waiting in a line without interacting.

Familiarity settings were highly consistent (inter-rater reliability: 0.97) across two independent coders suggestive of high accuracy as it is less likely for coders to be highly consistent in the way they are inaccurate (see Results). We confirmed this by having four participants validating familiarity judgements across the *entirety* of their own footage. Collectively, these participants validated a total of 830 familiarity judgements uncovering only 17 errors. This validation process conclusively showed that the coders were remarkably accurate in their familiarity judgements (98% correct), at least for the subset of four participants whose footage was submitted to validation.

In addition to these attributes, distances from which the faces were viewed were estimated. We mapped the size of faces on the digital images to distances from which they were captured by the camera in a two-step process. In step one, coders drew a line segment extending between two pose-dependent anchor points. For frontal and three-quarter faces, the anchor points were the two pupils; for profile faces, the anchor points were the tip of the nose to the tragus of the ear. In step two, we performed a size calibration procedure to estimate the transfer function of the camera that maps the physical distance of the face to the camera, to size of the line segment on the digital image. For this process, we recruited 10 female and 10 male models. We took photos of the models at seven known viewing distances (50 cm, 100 cm, 150 cm, 200 cm, 250 cm, 300 cm, 350 cm) and three poses (frontal, three-quarters, profile). We pooled photos based on gender and pose such that we obtained six groups of photos (3 poses × 2 genders). Each group of photos contained 70 images (10 models × 7 distances). These calibration photos were annotated by drawing line segments between the pose-dependent anchor points as described above in the same way as the participants' footage were annotated. For a given pose and gender, average line segment length in units of pixels at each known viewing distance was calculated by averaging across the measurements of the 10 models. A transfer function was obtained by fitting a parametric curve to these values. Six transfer functions were obtained, one for each of the gender-pose pairs in the following form $\log d = (\log l - a)/b$, where d denotes the viewing distance in cm, and l denotes the length of the line segment in pixels. Table 1 shows the values of the two parameters a and b for the six transfer functions.

Finally, a precise estimate of the viewing distance of each face detected in the participants' footage was obtained by converting the length of the annotated line segment using the appropriate transfer function obtained by the calibration process. Only adult faces were

Table 1
Size calibration fits. To estimate the relationship between viewing distance of each face based on face size on the digital image captured by our eyewear-embedded camera, line segments were drawn between pairs of pose-dependent anchor points on 10 female and 10 male faces at three poses (frontal, three-quarters, and profile) and at six known distances (50 cm, 100 cm, 150 cm, 200 cm, 250 cm, 300 cm, 350 cm). Parameters, a and b of the line fits obtained for the calibration data in form $\log d = (\log l - a)/b$, separately for six categories of faces (2 genders × 3 poses) are shown.

	a	b
Female, frontal	8.4	−0.86
Male, frontal	8.3	−0.83
Female, three-quarters	8.2	−0.86
Male, three-quarters	8.1	−0.84
Female, profile	8.7	−0.86
Male, profile	8	−0.72

annotated for size since the calibration procedure was based on adults and therefore would not produce accurate distance estimates for images of younger faces. Face sizes in degrees of visual angle were calculated directly from viewing distances based on a median face width of 12.8 cm for female faces and 14 cm for male faces (see Poston, 2000, page 72, face breadth statistics for median face width estimates). For size judgements, inter-rater reliability was defined as the correlation coefficient between the sizes of the line segments drawn by the coders on corresponding faces.

3. Results

3.1. Overall exposure

Out of a total of 25,195 frames, 4940 frames contained one or more faces. By individual participant, an average 22.5% (SD = 12.5%) of each participant's recorded frames contained one or more faces. In other words, participants were exposed to faces 13.5 min of every waking hour. An alternative estimate given by a weighted average taking into account overall length of the recording of each participant yielded a slightly lower estimate of 19.6%, suggesting just under 12 min of face exposure for each hour.

3.2. Face attributes

Out of a total 4940 frames that contained faces, 7641 faces were detected. By individual participant, an average of 254.7 (SD = 148.8) faces per participant were detected. The footage was analyzed twice in its entirety by independent coders. The reliability between the two analyses was high for all annotated attributes (all r's $>$ 0.96) except for frontal and three-quarters pose judgments where reliabilities were lower (r's = 0.85 and 0.76, respectively). All reliabilities are provided in Table 2. Student's t-tests carried out between frequency per participant for all pairs of attributes revealed no significant differences between the settings of the two coders, though p values of 0.08 and 0.07 were obtained for frontal and three-quarters pose frequencies, respectively, possibly suggesting a trend for uncertainty and/or disagreement regarding borderline viewpoints that lie between frontal and three-quarters poses. Inter-rater reliability was also high for face size annotations with correlations between sizes of corresponding line segments exceeding 99% for all 30 participants (all r's $>$ 0.99). In what follows (including the data figures) we report the results of one independent analysis. Results of the independent parallel analysis are similar and support the same conclusions. The data from both independent analyses are provided for direct comparison in Supplementary Table 1.

Table 2
Inter-rater reliability for all annotated face attributes. The entire footage obtained from 30 participants was independently analyzed twice. To assess the consistency between these two independent analyses, we estimated inter-rater reliability as the correlation between the frequency of each attribute (e.g. frequency of female faces) across the two independent analyses. The reliability of the face-size judgements were assessed via the correlation between the length of the line segments drawn on corresponding faces by the two coders per participant. All face-size reliabilities were uniformly greater than 0.99 across all 30 participants.

	Inter-rater reliability
Female	> 0.99
Male	> 0.99
Familiar	0.97
Unfamiliar	0.97
Caucasian	0.97
Asian	0.98
African	1
Frontal	0.85
Three-quarters	0.76
Profile	0.99
Size of line segment	> 0.99

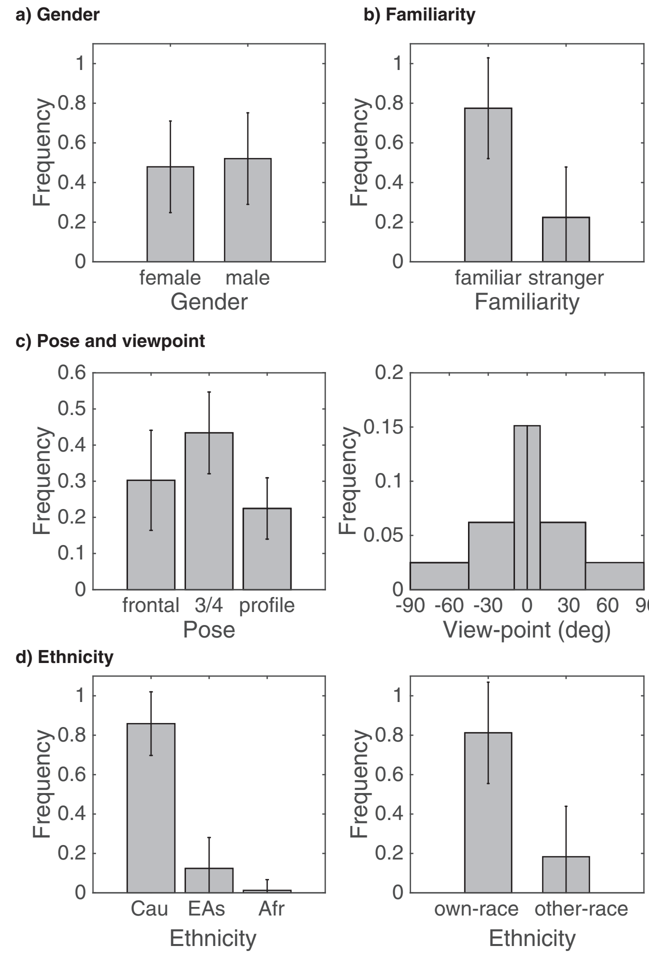

Fig. 2. Results. Exposure frequency averaged across participants for a) gender, b) familiarity, c) pose and viewpoint d) ethnicity. Error bars denote ± 1SD.

Fig. 2 shows detailed exposure statistics for gender, familiarity, pose and ethnicity. Results indicate that exposure to female faces (48%, SD = 23%) and male faces (52%, SD = 23%) did not differ significantly (p = .49). Exposure to familiar faces (77%, SD = 25%) exceeded exposure to unfamiliar faces (22%, SD = 25%) based on a two-tailed *t*-test ($p \ll .001$). Frequency of exposure to Caucasian faces (86%, SD = 16%) was significantly higher than exposure to both of the other two annotated ethnicities, Asian (12%, SD = 16%) and African (1%, SD = 5%), (both p's $\ll .001$), while exposure to Asian faces significantly exceeded exposure to African faces ($p < .01$). We also calculated frequencies for own- and other-race faces based on the 28 Caucasian participants (Fig. 2c, right panel). This analysis showed that Caucasian observers' own-race face exposure (81%, SD = 26%) significantly exceeded that of other-race faces (18%, SD = 26%) based on a two-sided *t*-test ($p \ll .001$). Based on the categorical definitions of frontal, three-quarters and profile poses (see Methods > Data analysis > Face annotation) majority of exposure were to faces in the three-quarters pose (44%, SD = 12%) which was significantly higher than exposure to faces in frontal (30%, SD = 14%) and profile (22%, SD = 8%) poses (both p's $\ll .001$), while exposure to frontal pose was significantly higher than that of profile pose ($p < .05$). We mapped our categorical pose definitions of 'frontal', 'three-quarters' and 'profile' to physical ranges of angles of head rotation based on a biometric database by Al Nizami et al. (2009). Assuming uniform distributions across these ranges, we computed adjusted frequencies per degrees of rotation angle by dividing total face frequency in a given pose category to the total rotation

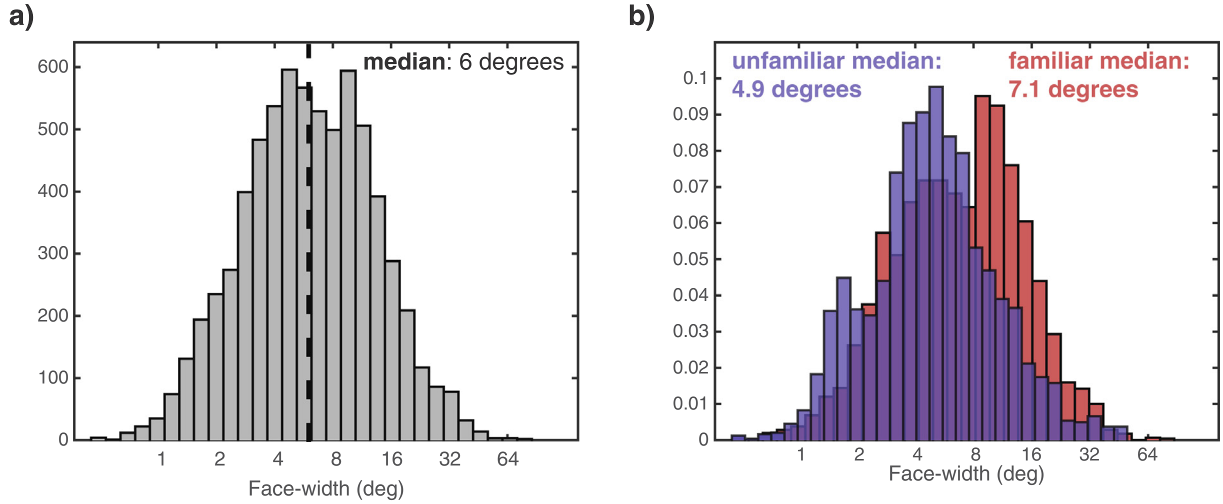

Fig. 3. Results. Distribution of face-widths a) across all faces b) across familiar and unfamiliar faces separately.

angle represented by that pose (Fig. 2d, right panel). For example, for frontal pose, the total frequency of 30% was distributed across the head rotation angle range of [−10° to 10°]. According to this representation, majority of views clustered around a frontal (i.e., 0°) pose.

Out of a total of 7641 detected faces, 7128 faces (93%) were annotated for size. Faces where pose was annotated as 'other' by one of the coders (i.e., poses that were not classified into one of the three calibrated poses) and those where one (or both) of the two pose-dependent anchor points were substantially occluded or out of the image frame, were not annotated for size. Distribution of estimated face sizes are shown in Fig. 3a. Median face size was 6°, corresponding to a median viewing distance of 128 cm. We hypothesized that face sizes may systematically depend on familiarity of the face. Indeed, face sizes were significantly larger for familiar (median = 7.1°) compared to unfamiliar (median = 4.9°) faces ($p \ll .01$). Fig. 3b shows the distribution of face sizes separately for familiar and unfamiliar faces corresponding to median viewing distances of 109 cm and 157 cm, respectively.

4. Discussion

We obtained first-person perspective images from thirty participants to examine the extent and statistics of daily face exposure in adults i.e., the adult face-diet. We aimed to determine whether and in what ways the adult face-diet differs from that of infants, and whether the contents of the adult face-diet can account for the various perceptual effects reported in the literature regarding adult face processing.

4.1. Overall face exposure

Our research revealed that faces have a prominent presence in the visual experience of adult observers. Based on a conservative estimate, adults spend 12 min of every waking hour exposed to faces. This degree of dense exposure to faces is matched only by the very early visual exposure of infants within the first few months after birth (Jayaraman et al., 2015; Sugden et al., 2014). Infants' face exposure declines systematically with age, reaching 5 min per hour by the end of 11 months, and remains at this level of exposure (or lower) at least until the end of 24 months of age (Fausey et al., 2016; Jayaraman et al., 2015, 2017). Why is adult face exposure substantially greater than that of infants older than 4 months of age? We know that an overwhelming majority of infants' face exposure is to familiar faces, most often of a family member, such as the mother (Jayaraman et al., 2015). Thus, one possibility is that the additional adult face exposure may be accounted for by encounters with strangers—a type of exposure that is largely missing from the infants' visual experience. When we limit our analysis of adult face exposure to familiar faces (i.e., exclude frames with no familiar faces, 2.82%) our overall exposure estimate reduces to just over 10 min of every hour. Therefore, even when exposure to familiar faces alone is considered, adult face exposure is approximately double the exposure of older infants.

4.2. Gender, ethnicity, and familiarity

No differences between exposure to female and male faces are found in the adult face-diet. This is aligned well with the demographics of the population that comprises approximately equal numbers of men and women. On the other hand, this pattern of exposure differs from that of infants who are exposed primarily to female faces (Sugden et al., 2014). Regarding ethnicity, adults are most frequently exposed to own-race faces. While this is consistent with infants' predominantly own-race face exposure (96%) (Sugden et al., 2014), unlike infants, adults are exposed to a non-negligible proportion of other-race faces (%18). Once again, one potential explanation is that the adults' other-race exposure may be mostly due to encounters with unfamiliar faces. When only familiar faces are considered, the own-race exposure of adult observers increase slightly, from 81% to 85%, yet, a non-negligible other-race exposure (14%) due to interactions with familiar individuals remains. Exposure to familiar faces doubles in adulthood compared to older infants. What is the source of this additional face exposure? We know that for the first year of an infant's life, primary exposure is to caregivers and family members, in particular, the mother (Jayaraman et al., 2015; Rennels & Davis, 2008). One possibility is that adults spend more time with family members compared to infants. However, given infants spend most of their time in the family home, engaging in home-based activities with family members (Sugden et al., 2014) and that adults engage in a wider variety of activities (e.g., occupation, physical activities, social activities) at a larger number of physical locations (e.g., home, work, community center, restaurant), we expect that adults' additional exposure to familiar faces likely comprise interactions with additional individuals, such as friends, acquaintances and colleagues. Although the sheer multitude of identities that were encountered in the adult participants' face-diets precluded us from carrying out a precise quantitative analysis of the distribution and number of unique identities, it is safe to assume that adults are exposed to a larger number of unique identities compared to infants. If development of face expertise requires exposure to a sufficient variety of unique identities, this may account for the late maturation of face recognition abilities in human observers.

4.3. Pose and viewpoint

Our results show that faces are encountered most commonly in the three-quarters pose in the adult face-diet. This differs from what has been reported for infants where majority of faces are viewed in the frontal pose. However, this comparison must be made with careful consideration to how different viewpoints (or poses) were defined by various studies. Sugden and Moulson (2017) have defined a face as 'frontal' if both eyes were visible and *"the face was pointing toward the camera"* and found 74% of faces were viewed in the frontal pose. Smith and colleagues defined 'frontal' as both eyes visible and found well over 50% face exposure to frontal faces (Jayaraman et al., 2015). Both of these definitions differ from our more restricted definition of frontal that included faces where both eyes and both ears would have been visible (if not for some occluder, such as hair). Based on a biometric database of head rotations (Al Nizami et al., 2009), we estimate that our frontal definition represents a range from a straight forward head position (i.e., 0°) up to about 10° head rotation in either direction, whereas both the infant studies' 'frontal' definitions would represent a larger range of rotations including up to a 20–30° rotation in either direction from directly frontal. When we adjust our percentages to conform to the frontal pose definitions of the infant studies by assuming half of the three-quarters faces in our study might qualify as frontal in the definition of the infant studies, we find that approximately 50% of the adult face exposure may be considered frontal. This rough estimate is consistent with that reported for infants by Jayaraman et al. (2015) but still considerably below the 74% estimate reported by Sugden and Moulson (2017). Since the infant studies did not explicitly code for three-quarters and profile poses, it is difficult to make any further comparisons regarding exposure to these poses.

Beyond the pose categories, one can examine frequency of exposure as a function of physical head-rotation angles. In principle, an isotropic frequency of exposure would be expected for an independent object that is free to rotate with respect to the viewer. In contrast, when adjusted for the physical ranges of head rotations represented by our three categorical pose definitions (based on Al Nizami et al., 2009), we find that the distribution is distinctly non-uniform, peaking around the central tendency of 0° (directly front facing) as seen in Fig. 2c, right panel. This is a clear suggestion that views of faces are fundamentally shaped by the nature of social interactions, which often place individuals facing each other, or nearly so. This argument is supported by a further analysis of viewpoint distribution for familiar faces only, which results in a lower frequency of profile pose (18% down from 22%) and higher frequency of frontal pose (33% up from 30%). Thus, when only familiar faces are considered, the view-point distribution deviated further from an isotropic pattern with an even higher peak around 0°, lending support to the argument that social interactions are a key factor in shaping exposure to faces in varying view-points.

View-point effects on face recognition have been studied extensively in the past and a so-called 'three-quarter pose advantage' has been reported by a number of studies (Krouse, 1981; O'Toole, Edelman, & Bulthoff, 1998; Troje & Bulthoff, 1996; Van der Linde & Watson, 2010), though not all, (e.g., see, Liu & Chaudhuri, 2002) which is most evident in the study phase, and where test and study faces differ in view-points (Troje & Bulthoff, 1996; Van der Linde & Watson, 2010). In other words, three-quarters is the optimal pose to learn a face for subsequent recognition at varying poses. Our results raise the possibility that the 'three quarters advantage' in face recognition may in fact arise as a consequence of visual experience with faces. Although our present results show that face views are encountered most commonly at 0°, the range of head rotations that generate prototypical 'frontal pose' face views are narrower than the range of head rotations that generate a prototypical 'three-quarters pose'. When face views are binned into these categories, maximal exposure occurs at the three-quarters view (Fig. 2c, left panel). Indeed, recognition performance reported in behavioural studies mirror the exposure statistics with intermediate performance at frontal pose and lowest performance at profile pose (Van der Linde & Watson, 2010). We argue that the view-point effects on face recognition performance may be a consequence of exposure statistics to faces.

4.4. Size

Our results show that in the adult face-diet, faces are relatively large (median 6°), i.e., viewed from fairly close distances (median 128 cm). These are generally in agreement with studies on the infant face-diet. Jayaraman et al. (2015) reported that mean viewing distances increased reliably with age from about 61 cm for the youngest infants to 122 cm for the oldest at 11 months of age. Sugden and Moulson (2017) reported that 87% of faces were viewed from 91 cm or closer. Both of these reports suggest that faces were somewhat larger (i.e., closer) in the face-diet of infants compared to adults. To estimate the contribution of unfamiliar faces in the adult face-diet to this difference in size, we re-analyzed the adult data for familiar faces only. This analysis showed that median size for familiar faces was 7.1°, i.e., viewed from a median distance of 109 cm, suggesting that face-size exposure is similar in adulthood and late infancy.

That larger faces are easier to recognize is accepted as common sense as well as demonstrated experimentally (Loftus & Harley, 2005; Lott, Haegerstrom-Portnoy, Schneck, & Brabyn, 2005). Yang, Shafai, and Oruc (2014) systematically examined face recognition efficiency as a function of face size and reported optimal processing beyond 6° face-width. The results of this latter study were consistent with better recognition at larger sizes (i.e. closer viewing distances), and additionally suggested a sharp switch in face processing strategies at the 6° size boundary (corresponding to about 128 cm viewing distance based on an average face-width of 13.4 cm, see biometric measurements of bizygomatic face breadth, Poston, 2000, page 72). Indeed, another study that examined critical spatial frequencies for face recognition as a function of size also suggested two distinct regimes of spatial frequency use that crossed over at a similar size boundary where larger faces were recognized using coarser features than predicted from smaller sizes (Oruc & Barton, 2010). McKone (2009) reported holistic processing to be strongest at viewing distances between 2 m and 10 m, falling off for closer viewing distances. These 'optimal distances' correspond to smaller faces, approximately 0.8°–3.8° face-width. Importantly, they found that larger faces showed weaker holistic processing. At face value these results are in contradiction with those of Yang et al. (2014). However, there was an important difference between the two studies: the McKone (2009) study did not assess recognition performance but rather studied measures that essentially represent face *detection*. Indeed, optimal viewing conditions for face detection might very well differ from those of recognition. Ross and Gauthier (2015) examined the magnitude of the composite-face effect, a measure of holistic processing that relates to recognition, as a function of face size. This study found holistic processing measures that grew stronger with larger face sizes, however, did not test faces larger than 4°, and therefore, does not offer further guidance in the present context.

Yang et al. (2014) reported evidence suggesting two distinct face recognition processes, one expert mechanism operating at face sizes larger than 6°, and another less efficient one, operating at 6° and smaller face sizes. The present study provides a potential explanation for this finding: these two viewing modes correspond well to the characteristic viewing distances to familiar (median 7.1°) vs. unfamiliar (median 4.9°) faces. Previous work has shown that experience with faces is the key factor in face expertise. Importantly, mere exposure without identification is not sufficient to develop expertise (Tanaka et al., 2013; Yovel et al., 2012). Thus, it is plausible that expert face processes are tuned to the characteristics of familiar faces (such as typical sizes of familiar faces). We propose that expert face processing is specifically adapted to the ecologically relevant viewing conditions for social interaction with familiar persons. Based on the present data, this tuning may occur within the lifetime of an observer. Alternatively, from an evolutionary perspective, exposure to strangers and unfamiliar individuals is a fairly recent development in human history, possibly at its height for individuals living in large metropolitan cities (such as the participants in our study). In contrast, ancestral experience may have been primarily based on familiar faces, and encounters with unfamiliar individuals may have been infrequent. Therefore, it is also conceivable that adaptation of face recognition processes to the viewing conditions that characterize social interactions with familiar individuals is a process that has emerged in evolutionary timescales.

4.5. Future directions

Overall, our results show that characteristics of experience with faces can account for prominent behavioural findings regarding face recognition, such as the other-race effect and the three-quarters pose advantage. Future studies are required to examine more subtle findings, such as the own-gender effect, exploring the possibility that the reduced reliability in the observation of such effects may stem from greater individual differences in the contents of the face-diet for such attributes, such as face gender. Also of particular interest would be studies to examine face-diet in specific sub-populations, such as differences between females and males and in clinical populations, such as the autism spectrum disorder population.

5. Conclusions

Adults' face exposure is denser, richer, and more diverse than that of infants: Adults visually experience familiar faces approximately twice as long as infants and view a larger variety of faces. This is likely due to the greater variety of activities adults take part in. Sugden et al. (2014) has raised the possibility that the restricted range and homogeneity of faces seen in infancy may play an important role in the visual development and perceptual narrowing regarding face recognition. We speculate that the gradual enrichment of face experience with age may be a key factor underlying protracted development of face recognition abilities in human observers. Further, we argue that experience with familiar faces shape face recognition processes by tuning them to the image properties characterized by the context of social interaction. This includes tuning for face sizes characteristic of close-viewing conditions. Whether the effect of this experience is brought on through evolutionary time scales, during sensitive periods of development, or through the lifetime of an individual remains unclear.

Acknowledgments

This work was supported by a Natural Sciences and Engineering Research Council of Canada, Discovery Grant RGPIN 402654-11 (IO) and a Canada Foundation for Innovation, John R. Evans Leaders Fund (IO). We thank Ramin Hamidizadeh for his help and recommendation regarding the choice of camera and with the preparation of the research ethics application, and Aliana Carlos for her assistance in coding the video frames.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.visres.2018.01.001.

Reference

Al Nizami, H. A., Adkins-Hill, J. P., Zhang, Y., Sullins, J. R., McCullough, C., Canavan, S., & Yin, L. (2009). *A biometric database with rotating head videos and hand-drawn face sketches*. Paper presented at the Biometrics: Theory, Applications, and Systems, 2009. BTAS'09. IEEE 3rd International Conference on.

Carey, S., Diamond, R., & Woods, B. (1980). Development of Face Recognition—A Maturational Component? *Developmental Psychology, 16*(4), 257–269.

de Heering, A., & Maurer, D. (2014). Face memory deficits in patients deprived of early visual input by bilateral congenital cataracts. *Developmental Psychobiology, 56*(1), 96–108. http://dx.doi.org/10.1002/dev.21094.

de Heering, A., Rossion, B., & Maurer, D. (2012). Developmental changes in face recognition during childhood: Evidence from upright and inverted faces. *Cognitive Development, 27*, 17–27.

Farah, M. J., Rabinowitz, C., Quinn, G. E., & Liu, G. T. (2000). Early commitment of neural substrates for face recognition. *Cognitive Neuropsychology, 17*(1), 117–123. http://dx.doi.org/10.1080/026432900380526.

Fausey, C. M., Jayaraman, S., & Smith, L. B. (2016). From faces to hands: Changing visual input in the first two years. *Cognition, 152*, 101–107. http://dx.doi.org/10.1016/j.cognition.2016.03.005.

Fine, I., Wade, A. R., Brewer, A. A., May, M. G., Goodman, D. F., Boynton, G. M., ... MacLeod, D. I. (2003). Long-term deprivation affects visual perception and cortex. *Nature Neuroscience, 6*(9), 915–916. http://dx.doi.org/10.1038/nn1102.

Geldart, S., Mondloch, C. J., Maurer, D., De Schonen, S., & Brent, H. P. (2002). The effect of early visual deprivation on the development of face processing. *Developmental Science, 5*(4), 490–501.

Germine, L. T., Duchaine, B., & Nakayama, K. (2011). Where cognitive development and aging meet: face learning ability peaks after age 30. *Cognition, 118*(2), 201–210. http://dx.doi.org/10.1016/j.cognition.2010.11.002.

Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics, 56*(4), 544–549. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1165958.

Hayward, W. G., Rhodes, G., & Schwaninger, A. (2008). An own-race advantage for components as well as configurations in face recognition. *Cognition, 106*(2), 1017–1027. http://dx.doi.org/10.1016/j.cognition.2007.04.002.

Huber, E., Webster, J. M., Brewer, A. A., MacLeod, D. I., Wandell, B. A., Boynton, G. M., ... Fine, I. (2015). A lack of experience-dependent plasticity after more than a decade of recovered sight. *Psychological Science, 26*(4), 393–401. http://dx.doi.org/10.1177/0956797614563957.

Jayaraman, S., Fausey, C. M., & Smith, L. B. (2015). The faces in infant-perspective scenes change over the first year of life. *PLoS ONE, 10*(5), e0123780. http://dx.doi.org/10.1371/journal.pone.0123780.

Jayaraman, S., Fausey, C. M., & Smith, L. B. (2017). Why are faces denser in the visual experiences of younger than older infants? *Developmental Psychology, 53*(1), 38–49. http://dx.doi.org/10.1037/dev0000230.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40*(1–2), 1–19. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1786670.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: evidence of perceptual narrowing. *Psychological Science, 18*(12), 1084–1089. http://dx.doi.org/10.1111/j.1467-9280.2007.02029.x.

Krouse, F. L. (1981). Effects of pose, pose change, and delay on face recognition performance. *Journal of Applied Psychology, 66*(5), 651.

Laguesse, R., Dormal, G., Biervoye, A., Kuefner, D., & Rossion, B. (2012). Extensive visual training in adulthood significantly reduces the face inversion effect. *Journal of Vision, 12*(10), 14. http://dx.doi.org/10.1167/12.10.14.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Neuroperception. Early visual experience and face processing. *Nature, 410*(6831), 890. http://dx.doi.org/10.1038/35073749.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2003). Expert face processing requires visual input to the right hemisphere during infancy. *Nature Neuroscience, 6*(10), 1108–1112. http://dx.doi.org/10.1038/nn1121.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2004). Impairment in holistic face processing following early visual deprivation. *Psychological Science, 15*(11), 762–768. http://dx.doi.org/10.1111/j.0956-7976.2004.00753.x.

Lewis, T. L., & Maurer, D. (2005). Multiple sensitive periods in human visual development: evidence from visually deprived children. *Developmental Psychobiology, 46*(3), 163–183. http://dx.doi.org/10.1002/dev.20055.

Liu, C. H., & Chaudhuri, A. (2002). Reassessing the 3/4 view effect in face recognition. *Cognition, 83*(1), 31–48.

Loftus, G. R., & Harley, E. M. (2005). Why is it easier to identify someone close than far away? *Psychonomic Bulletin & Review, 12*(1), 43–65. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15948283.

Lott, L. A., Haegerstrom-Portnoy, G., Schneck, M. E., & Brabyn, J. A. (2005). Face recognition in the elderly. *Optometry & Vision Science, 82*(10), 874–881.

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

McKone, E., Crookes, K., Jeffery, L., & Dilks, D. D. (2012). A critical review of the development of face recognition: experience is less important than previously believed. *Cognitive Neuropsychology, 29*(1–2), 174–212. http://dx.doi.org/10.1080/02643294.2012.660138.

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences, 11*(1), 8–15. http://dx.doi.org/10.1016/j.tics.2006.11.002.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7*(1), 3–35. http://dx.doi.org/10.1037/1076-8971.7.1.3.

O'Toole, A. J., Deffenbacher, K. A., Valentin, D., & Abdi, H. (1994). Structural aspects of face recognition and the other-race effect. *Memory & Cognition, 22*(2), 208–224. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8035697.

O'Toole, A. J., Edelman, S., & Bulthoff, H. H. (1998). Stimulus-specific effects in face recognition over changes in viewpoint. *Vision Research, 38*(15–16), 2351–2363. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9798004.

Oruc, I., & Barton, J. J. (2010). Critical frequencies in the perception of letters, faces, and novel shapes: evidence for limited scale invariance for faces. *Journal of Vision, 10*(12), 20. http://dx.doi.org/10.1167/10.12.20.

Poston, A. (2000). *Human Engineering Design Data Digest*. Washington, DC Retrieved from http://hfetag.com/hfs_docs.html.

Rennels, J. L., & Davis, R. E. (2008). Facial experience during the first year. *Infant Behavior and Development, 31*(4), 665–678. http://dx.doi.org/10.1016/j.infbeh.2008.04.009.

Rhodes, G., Hayward, W. G., & Winkler, C. (2006). Expert face coding: configural and component coding of own-race and other-race faces. *Psychonomic Bulletin & Review, 13*(3), 499–505. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17048737.

Rhodes, G., Jeffery, L., Watson, T. L., Clifford, C. W., & Nakayama, K. (2003). Fitting the mind to the world: face adaptation and attractiveness aftereffects. *Psychological Science, 14*(6), 558–566. http://dx.doi.org/10.1046/j.0956-7976.2003.psci_1465.x.

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

Rostamirad, S., Barton, J. J., & Oruc, I. (2009). Center-surround organization of face-space: evidence from contrast-based face-priming. *NeuroReport, 20*(13), 1177–1182. http://dx.doi.org/10.1097/WNR.0b013e32832f078d.

Sangrigoli, S., Pallier, C., Argenti, A. M., Ventureyra, V. A., & de Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science, 16*(6), 440–444. http://dx.doi.org/10.1111/j.0956-7976.2005.01554.x.

Sugden, N. A., Mohamed-Ali, M. I., & Moulson, M. C. (2014). I spy with my little eye: typical, daily exposure to faces documented from a first-person infant perspective. *Developmental Psychobiology, 56*(2), 249–261. http://dx.doi.org/10.1002/dev.21183.

Sugden, N. A., & Moulson, M. C. (2017). Hey baby, what's "up"? One- and 3-month-olds experience faces primarily upright but non-upright faces offer the best views. *Quarterly Journal of Experimental Psychology (Hove), 70*(5), 959–969. http://dx.doi.org/10.1080/17470218.2016.1154581.

Tanaka, J. W., Heptonstall, B., & Hagen, S. (2013). Perceptual expertise and the plasticity of other-race face recognition. *Visual Cognition, 21*, 1183–1201.

Troje, N. F., & Bulthoff, H. H. (1996). Face recognition under varying poses: the role of texture and shape. *Vision Research, 36*(12), 1761–1771. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8759445.

Van der Linde, I., & Watson, T. (2010). A combinatorial study of pose effects in unfamiliar face recognition. *Vision Research, 50*(5), 522–533. http://dx.doi.org/10.1016/j.visres.2009.12.012.

Walker, P. M., & Tanaka, J. W. (2003). An encoding advantage for own-race versus other-race faces. *Perception, 32*(9), 1117–1125. http://dx.doi.org/10.1068/p5098.

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., ... Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of Sciences, 107*(11), 5238–5241. http://dx.doi.org/10.1073/pnas.0913053107.

Yang, N., Shafai, F., & Oruc, I. (2014). Size determines whether specialized expert processes are engaged for recognition of faces. *Journal of Vision, 14*(8), 17. http://dx.doi.org/10.1167/14.8.17.

Yovel, G., Halsband, K., Pelleg, M., Farkash, N., Gal, B., & Goshen-Gottstein, Y. (2012). Can massive but passive exposure to faces contribute to face recognition abilities? *Journal of Experimental Psychology: Human Perception and Performance, 38*(2), 285–289. http://dx.doi.org/10.1037/a0027077.

Zhu, Q., Song, Y., Hu, S., Li, X., Tian, M., Zhen, Z., ... Liu, J. (2010). Heritability of the specific cognitive ability of face perception. *Current Biology, 20*(2), 137–142. http://dx.doi.org/10.1016/j.cub.2009.11.067.