



## Page Proof Instructions and Queries

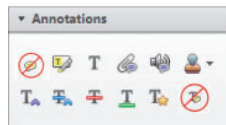
**Journal Title:** Perception (PEC)

**Article Number:** 652044

Greetings, and thank you for publishing with SAGE. We have prepared this page proof for your review. Please respond to each of the below queries by digitally marking this PDF using Adobe Reader.

Click "Comment" in the upper right corner of Adobe Reader to access the mark-up tools as follows:

For textual edits, please use the "Annotations" tools. Please refrain from using the two tools crossed out below, as data loss can occur when using these tools.



For formatting requests, questions, or other complicated changes, please insert a comment using "Drawing Markups."



Detailed annotation guidelines can be viewed at: <http://www.sagepub.com/repository/binaries/pdfs/AnnotationGuidelines.pdf>

Adobe Reader can be downloaded (free) at: <http://www.adobe.com/products/reader.html>.

No.	Query
	Please confirm that all author information, including names, affiliations, sequence, and contact details, is correct.
	Please review the entire document for typographical errors, mathematical errors, and any other necessary corrections; check headings, tables, and figures.
	Please confirm that the Funding and Conflict of Interest statements are accurate.
	Please ensure that you have obtained and enclosed all necessary permissions for the reproduction of artistic works, (e.g. illustrations, photographs, charts, maps, other visual material, etc.) not owned by yourself. Please refer to your publishing agreement for further information.
	Please note that this proof represents your final opportunity to review your article prior to publication, so please do send all of your changes now.
AQ: 1	Per style, abstracts should not have reference citations and hence the citation "Short & Mondloch, 2013" has been deleted from the abstract. Kindly confirm.
AQ: 2	Please provide editor(s) name for 'Rhodes et al. (2005)'.

# Judging Normality and Attractiveness in Faces: Direct Evidence of a More Refined Representation for Own-Race, Young Adult Faces

*Perception*

2016, 0(0) 1–18

© The Author(s) 2016

Reprints and permissions:

[sagepub.co.uk/journalsPermissions.nav](http://sagepub.co.uk/journalsPermissions.nav)

DOI: 10.1177/0301006616652044

[pec.sagepub.com](http://pec.sagepub.com)



**Xiaomei Zhou**

Department of Psychology, Brock University, St Catharines, ON, Canada

**Lindsey A. Short**

Department of Psychology, Brock University, St Catharines, Canada;  
Department of Psychology, Redeemer University College, Ancaster,  
Canada

**Harmonie S. J. Chan**

Department of Psychology, Brock University, St Catharines, Canada

**Catherine J. Mondloch**

Department of Psychology, Brock University, St Catharines, Canada;  
ARC Centre of Excellence in Cognition and Its Disorders, University of  
Western Australia, Perth, Australia

## Abstract

Young and older adults are more sensitive to deviations from normality in young than older adult faces, suggesting that the dimensions of face space are optimized for young adult faces. Here, we extend these findings to own-race faces and provide converging evidence using an attractiveness rating task. In Experiment 1, Caucasian and Chinese adults were shown own- and other-race face pairs; one member was undistorted and the other had compressed or expanded features. Participants indicated which member of each pair was more normal (a task that requires referencing a norm) and which was more expanded (a task that simply requires discrimination). Participants showed an own-race advantage in the normality task but not the discrimination task. In Experiment 2, participants rated the facial attractiveness of own- and other-race faces (Experiment 2a) or young and older adult faces (Experiment 2b). Between-rater variability in ratings of individual faces was higher for other-race and older adult faces; reduced consensus in attractiveness judgments reflects a less refined face space. Collectively, these results provide direct evidence that the dimensions of face space are optimized for own-race and young adult faces, which may underlie face race- and age-based deficits in recognition. **AQ1**

---

## Corresponding author:

Catherine J. Mondloch, Department of Psychology, Brock University, 500 Glenridge Avenue, St. Catharines, ON, Canada, L2S 3A1.

Email: [cmondloch@brocku.ca](mailto:cmondloch@brocku.ca)

**Keywords**

other-race effect, young adult recognition bias, perceptual expertise, face space

Adults possess an exceptional ability to discriminate and recognize individual faces, despite the fact that all faces share the same configural template (i.e., two eyes located above the nose and mouth). This perceptual expertise has been attributed to norm-based coding, a process by which individual face exemplars are encoded with reference to their deviation from the face norm (i.e., center of face space), which represents the average of all faces previously encountered (Valentine, 1991). Strong evidence for norm-based coding has emerged from studies examining face aftereffects (e.g., Leopold, O'Toole, Vetter, & Blanz, 2001; Rhodes & Jeffery, 2006; Rhodes et al., 2005; Schweinberger et al., 2010). For example, repeated exposure to an adaptor face (e.g., anti-Dan) shifts the norm toward that face, biasing perception selectively toward a face with attributes opposite to the adaptor (e.g., Dan; termed identity aftereffects). Likewise, exposure to faces distorted in a similar direction (e.g., features expanded outward) produces a temporary shift in the norm, such that unaltered faces appear distorted in the opposite direction while similarly distorted faces appear more attractive (termed figural aftereffects; Rhodes, Jeffery, Watson, Clifford, & Nakayama, 2003).

Norm-based coding is functionally important; it allows efficient extraction of subtle variations in the shared configuration among faces (Byatt & Rhodes, 1998). This process frees up neural resources by allowing the perceptual system to focus on the unique characteristics that are crucial for identifying a particular face, rather than storing a complete structural description of each face (Rhodes & Leopold, 2011; Rhodes, Watson, Jeffery, & Clifford, 2010; Webster & MacLeod, 2011). The functional value of norm-based coding is evident in the positive correlation between the magnitude of both figural eye-height (Dennett, McKone, Edwards, & Susilo, 2012) and identity aftereffects (Rhodes, Jeffery, Taylor, Hayward & Ewing, 2014) and individual differences in face recognition memory.

Valentine's (1991) influential norm-based coding model provides an elegant explanation for two well-known phenomena: the other-race effect (ORE; better recognition of own- than other race faces; see Bothwell, Brigham & Malpass, 1989; Meissner & Brigham, 2001 for reviews) and the other-age effect (OAE; better recognition of own-age; Anastasi & Rhodes, 2005; Perfect & Harris, 2003; Rhodes & Anastasi, 2012) or young adult faces (de Heering & Rossion, 2008; Kuefner, Macchi Cassia, Picozzi, & Bricolo, 2008) than faces from other age categories. Valentine proposed that the dimensions of face space are shaped by experience such that they maximally differentiate faces from categories with which adults have abundant experience. Consequently, faces from other categories (typically other-race and other-age faces) are densely clustered in the periphery of face space, making them hard to discriminate and recognize.

Category-contingent aftereffects provide partial support for this explanation. Adaptation to face categories distorted in opposite directions (e.g., compressed Caucasian versus expanded Asian faces) simultaneously shifts normality/attractiveness preferences in opposite directions. Opposing aftereffects have been found for race (Jaquet, Rhodes, & Hayward, 2008; Little, DeBruine, Jones, & Waitt, 2008), orientation (Rhodes et al., 2004), species (Little et al., 2008), sex (Jaquet & Rhodes, 2008; Little, DeBruine, & Jones, 2005), and age (Little et al., 2008; Short, Proietti, & Mondloch, 2015) and suggest that separable norms are used to code faces from different face categories (e.g., race and species).

Although opposing aftereffects demonstrate that adults possess separable norms coding for faces from different categories, they do not address whether the norm and face space are

less well differentiated for categories with which adults have less experience (e.g., other-race faces), a critical component of Valentine's model. This is because the magnitude of aftereffects does not vary as a function of perceptual expertise; adults do not show larger aftereffects for upright faces compared to inverted faces (Rhodes et al., 2004), for own-race compared to other-race faces (Jaquet et al., 2008), or for young adult compared to older adult faces (Short et al., 2015). Thus, the refinement of the norms and underlying dimensions used for faces from different categories remains unclear. In the current study, we directly test whether adults possess a more refined face space for own-race and young adult faces relative to other-race and older adult faces, respectively.

The method used to address this question in Experiment 1 is based on a previous study showing that poor recognition of older faces may be partially attributable to insensitivity to deviations from the norm in older relative to young adult faces (Short & Mondloch, 2013). In that study, participants were shown young and older adult face pairs in which one face was undistorted and the other image of the same identity had either compressed or expanded facial features. Participants were asked to judge which face in each face pair was more normal (normality task) and which was more expanded (discrimination task). The normality task is sensitive to norm-based coding because participants need to reference a norm. In contrast, the discrimination task does not require referencing a norm because participants can simply compare feature size. Short and Mondloch found that both young and older adults were more accurate for young than older adult faces in the normality task whereas they exhibited comparable accuracy for young and older faces in the discrimination task. Enhanced performance for young adult faces in the normality task was presumed to reflect reliance on a face space optimized for the dimensions of young adult faces, perhaps due to the early and continuous exposure most adults have to young adult faces throughout the lifespan. Consistent with this viewpoint, a recent study demonstrated that this enhanced sensitivity to the dimensions of young relative to older faces emerges as early as 3 years of age (Short, Mondloch, & Hackland, 2015).

If abundant experience with young adult faces sets up the perceptual system in a way that is preferentially tuned for the dimensions of young adult faces, then we would expect that abundant experience with own-race faces tunes the dimensions of face space for own-race faces in the same way, resulting in enhanced normality judgments for own- relative to other-race faces. In the current study, we directly tested this hypothesis. In Experiment 1, Caucasian and Chinese adults were tested with a modified version of the normality and discrimination tasks employed by Short and Mondloch (2013). Participants viewed own- and other-race face pairs rather than young and older adult face pairs. We predicted that both Caucasian and Chinese adults would be more accurate in judging the normality of own-race faces but would show comparable accuracy in discriminating own- and other-race faces.

In Experiment 2, to provide converging evidence for the conclusion that the dimensions of face space are more refined for the face categories with which adults have ample experience, we tested adults' sensitivity to variability in the attractiveness of own- versus other-race faces (Experiment 2a) and young versus older adult faces (Experiment 2b). Considerable evidence has suggested that similar to perceptions of normality, adults' perception of attractiveness is influenced by norm-based coding (e.g., Langlois & Roggman, 1990; Rhodes, 2006; Rhodes & Tremewan, 1996). Perceived facial attractiveness changes as a function of proximity to the norm, such that perceptual attractiveness decreases the farther a face is from the norm (e.g., Langlois & Roggman, 1990). If our mental representation of young, own-race faces is characterized by a more well-refined face space, then young, own-race faces should be more dispersed than older and other-race faces (Burton & Vokey, 1998), increasing consensus not only in judgments of normality but also attractiveness.

If abundant experience with own-race and young adult faces makes the dimensions of face space more finely tuned for these face categories, then adults should be more likely to agree on (i.e., greater consensus; less between-rater variability) the attractiveness of individual faces from such face categories relative to face categories with which they have less experience. To test this hypothesis, we presented participants with undistorted images of own- and other-race faces (Experiment 2a) or young and older adult faces (Experiment 2b) and measured the extent to which participants agreed on the attractiveness of each face (between-rater variability), which is quantified by the magnitude of the standard deviation (SD) in attractiveness ratings for each face. We hypothesized that participants would show greater agreement (i.e., smaller mean standard deviations) for face categories with which they had greater cumulative life experience (i.e., own-race and young adult faces) than faces with which they had limited perceptual experience (i.e., other-race and older adult faces).

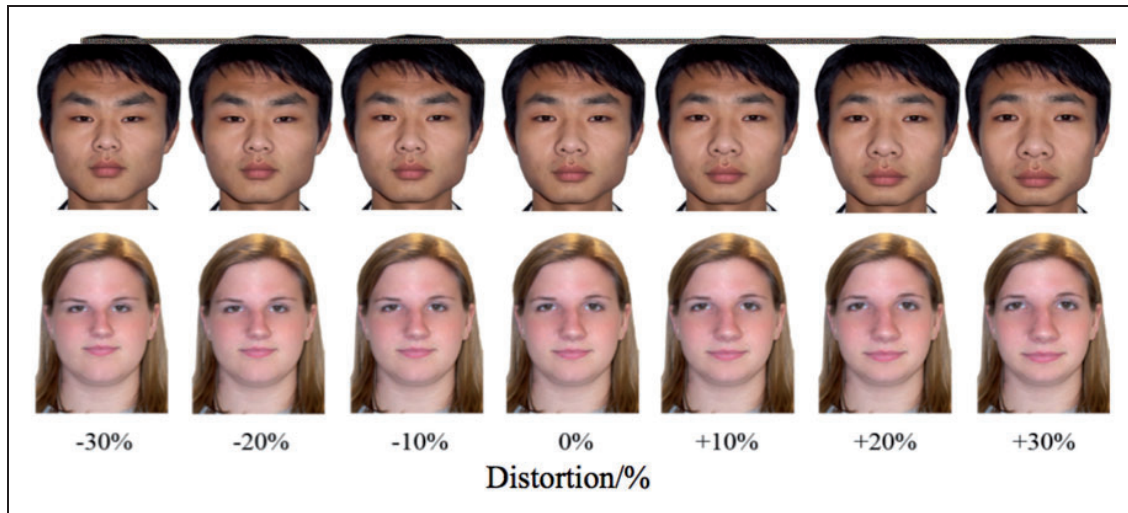
## Experiment 1

### Method

**Participants.** Twenty-four Caucasian adults (24 female; mean age = 19.67 years,  $SD = 1.17$ , age range = 18–22) from Brock University, Canada and 24 Chinese adults (18 female; mean age = 22.00 years,  $SD = 2.25$ , age range = 20–27) from Zhejiang Normal University, China participated in this experiment. All participants included in our analyses reported minimal contact with other-race identities based on their responses on a questionnaire (see Procedure section). All participants reported having less than two other-race friends and 98.75% of participants reported having zero other-race friends. We excluded five additional participants (three Caucasian and two Chinese adults) who reported significant experience with individuals of East Asian/Caucasian ethnicity. All participants gave written informed consent and received either research credit or a small honorarium for their participation.

**Materials.** Stimuli consisted of colored facial photographs of 12 Caucasian adults (6 male) and 12 Chinese adults (6 male). All stimuli were acquired from the Center for Vital Longevity Face Database (Minear & Park, 2004) and the Let's Face It database at Brock University. Faces were presented in a frontal view and posed a neutral expression. Faces were resized such that the distance from the hairline to the chin was approximately 450 pixels, and the spherize tool in Adobe Photoshop Version CS5 was used to expand and compress the internal features of each face (see Figure 1). The facial features of each face were either expanded outward or compressed inward at three distortion levels (–30%, –20%, –10%, 10%, 20%, 30%), resulting in a total of six new versions of each face. Each level of distortion was then paired with the undistorted same-identity counterpart. The left/right position of the undistorted face in each face pair was counterbalanced such that the undistorted face appeared on the left side in half of the trials. An additional four identities were used in four practice trials. Practice trials consisted of an undistorted face paired with either an expanded or compressed face ( $\pm 40\%$ ) of the same identity. The practice and test stimuli were approximately  $33 \times 20$  cm when presented on a 23-in. computer monitor and were viewed from a distance of approximately 60 cm. Stimuli were presented and participants' responses were recorded using SuperLab 4.5 software.

**Procedure.** This study received clearance from the Research Ethics Board at Brock University. All participants were tested individually in two tasks: a normality judgment task and a discrimination task. The order in which participants completed the two tasks was counterbalanced such that half of the participants were tested with the normality task



**Figure 1.** Sample distortion continua for a Chinese identity and a Caucasian identity. Each face pair comprises an undistorted face paired with a compressed or an expanded version of the same identity.

followed by the discrimination task and the other half were tested in the reverse order. In both tasks, each trial comprises a 500-ms fixation cross, followed by a face pair that was presented for 3,000 ms. Once the face pair disappeared from the screen, it was replaced by a screen prompting participants to press a key indicating which face was either more normal-looking (normality task) or more expanded (discrimination task). Within each task, race of face was blocked; half of the participants were tested with Caucasian faces first and the other half were tested with East Asian faces first. Within each block, both face identities and distortion levels were randomized for each participant. In the discrimination task, the expanded face in each pair was defined as having more stretched (i.e., expanded outward) features than its same-identity counterpart (e.g., undistorted face compared to a  $-10\%$  face). Twelve identities across six levels of distortion for each of the two race categories resulted in a total of 144 face pairs that were presented to participants in each task. Prior to the actual test, participants were presented with four practice trials in which the distortion was increased to  $\pm 40\%$ . On each practice trial, participants were shown an undistorted face paired with an expanded or compressed face of the same identity.

Upon completion of both tasks, participants completed a questionnaire assessing the amount of contact they have had with other-race identities (e.g., Chinese participants' contact with Caucasian individuals). For example, they indicated how many of their top 10 friends were of East Asian/Caucasian ethnicity, and how much current and previous experience they have had with other-race identities.

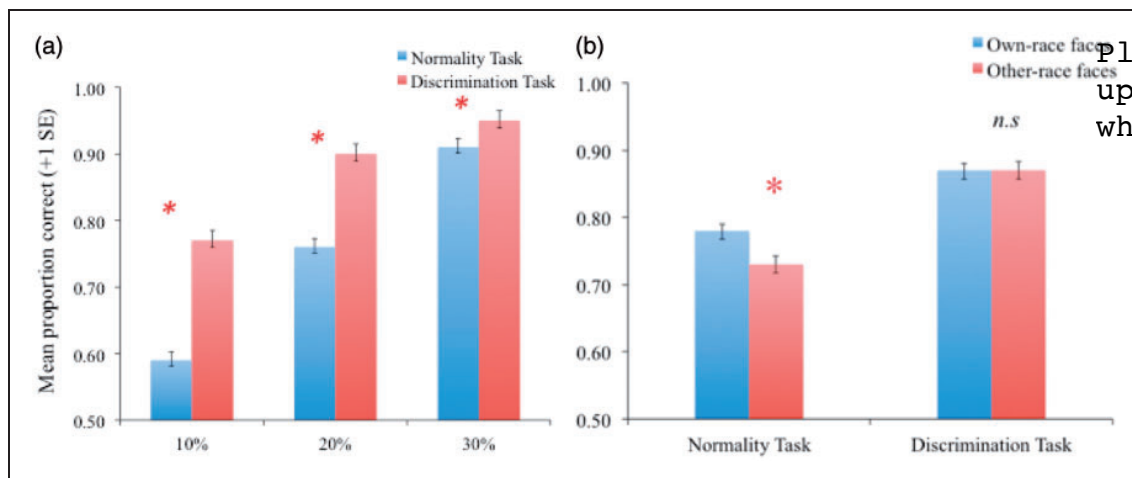
### Results and Discussion

To simplify our analysis, we collapsed across expanded and compressed trials within each distortion level. In the normality task, we calculated the proportion of trials in which participants selected the undistorted face in a face pair as being more normal. In the discrimination task, we calculated the proportion of trials in which participants selected the more expanded face in a face pair as being more expanded (i.e., the undistorted face when the distorted face was compressed, but the expanded face when the distorted face was expanded). Task order did not influence the accuracy of normality judgments and discrimination ( $ps > .12$ ); thus, task order was excluded from all further analyses.

We conducted a 2 (participant race: Caucasian, Chinese)  $\times$  2 (task type: normality, discrimination)  $\times$  2 (face race: Caucasian, East Asian)  $\times$  3 (distortion: 10%, 20%, 30%) **mixedmodel** analysis of variance (ANOVA), with participant race as a between-subjects variable and task type, face race, and distortion levels as within-subjects variables.

Because our primary question concerned the influence of face race on normality versus discrimination judgments, we focus here on main effects and interactions involving task type. We found significant main effects of task type,  $F(1, 46) = 60.86, p < .001, \eta_p^2 = .57$ ; face race,  $F(1, 46) = 21.99, p < .001, \eta_p^2 = .32$ ; and distortion level,  $F(1, 45) = 540.55, p < .001, \eta_p^2 = .96$ , such that accuracy was higher in the discrimination task ( $M = 0.87, SE = 0.01$ ) than in the normality task ( $M = 0.76, SE = 0.01$ ) and for own-race faces ( $M = 0.83, SE = 0.01$ ) than for other-race faces ( $M = 0.80, SE = 0.01$ ). Accuracy increased as distortion level increased (10%:  $M = 0.68, SE = 0.01$ ; 20%:  $M = 0.83, SE = 0.01$ ; 30%:  $M = 0.93, SE = 0.01$ ). There was also a task type by distortion interaction,  $F(1, 45) = 57.68, p < .001, \eta_p^2 = .72$ , which indicated that the difference in performance across the two tasks decreased as distortion level increased (i.e., the task became easier; see Figure 2(a)).

Notably, we found a significant interaction between task type and face race,  $F(1, 46) = 15.00, p < .001, \eta_p^2 = .25$  (see Figure 2(b)). Paired-sample  $t$  tests confirmed that both Caucasian and Chinese participants were more accurate in judging the normality of own-race faces ( $M = 0.78, SE = 0.01$ ) than other-race faces ( $M = 0.73, SE = 0.01, t_{47} = 5.42, p < .001, \text{Cohen's } d = 0.59$ ). In contrast, accuracy for own- ( $M = 0.87, SE = 0.01$ ) and other-race ( $M = 0.87, SE = 0.01$ ) faces did not differ in the discrimination task ( $t_{47} = 5.34, p = 0.60, \text{Cohen's } d = 0$ ). This task type by face race interaction was not qualified by interactions with either participant race or distortion level,  $ps > 0.41$ . Furthermore, when examined independently, both Caucasian and Chinese participants showed an own-race advantage on the normality task ( $ps < .010$ ) but not the discrimination task ( $ps > .100$ ). To determine whether participants' greater accuracy in judging the normality of own-race faces than other-race faces is attributable to better image learning **in** the course of the experiment for own- than other-race faces (i.e., to participants recognizing previously selected faces), we conducted an ANOVA in which block (first **vs** second half of trials), face race, participant race, and distortion levels were factors. There was no effect of block and no interaction involving block was significant ( $ps > .325$ ). Thus, more accurate performance for



**Figure 2.** Mean proportion correct for the discrimination and normality tasks at (a) each level of distortion collapsed across face race and for (b) own- and other-race faces collapsed across distortion levels. \* $p < .05$ .

Please use the updated black and white figure here

own- than for other-race faces in the normality task is not attributable to image learning (i.e., greater sensitivity to normality in own-race faces did not emerge **overtime**).

In summary, both Caucasian and Chinese adults showed deficits in detecting the normality of other-race compared to own-race faces despite no effect of face race on their ability to detect the expansion of facial features. This is consistent with evidence that young and older adults show a young adult face advantage in the normality but not the discrimination task (Short & Mondloch, 2013). Discrepant results across the two tasks directly points to reduced efficiency in the use of norm-based coding for other-race faces, another category with which most people have less experience. Extensive perceptual experience with own-race faces tunes the dimensions of face space for own-race faces, making judgments of normality more accurate for own- than other-race faces. But lack of perceptual experience with other-race faces does not influence accuracy in the discrimination task, a task that does not require the use of norm-based coding. This argument receives support from evidence that inverting the face pairs eliminates the young adult face advantage in the normality task but not the discrimination task (Short & Mondloch, 2013).

## Experiment 2

Similar to judgments of normality, judgments of attractiveness are influenced in part by how much an individual face deviates from an average face; facial attractiveness is inversely related to distance from the mean (Langlois & Roggman, 1990; Light, Hollander, & Kayra-Stuart, 1981; Morris & Wickham, 2001; O'Toole, Deffenbacher, Valentin, & Abdi, 1994; Rhodes, 2006; Rhodes & Tremewan, 1996; but see Alley & Cunningham, 1991). Similar mean attractiveness ratings across different cultures (Cunningham, Roberts, Barbee, Druen, & Wu, 1995; Langlois et al., 2000) and age groups (Cross & Cross, 1971) suggests a degree of consensus regarding which faces are most versus least attractive. However, mean ratings for an individual face ignore between-rater variability (consensus), a metric that we hypothesized would be influenced by experience.

If extensive experience with own-race and young adult faces optimizes the dimensions of face space for these face categories, resulting in faces from other categories being more densely clustered with poorly refined norms, then there should be less consensus among raters when judging faces from categories with which they have less experience. To test this hypothesis and provide converging evidence for a face space optimized for own-race and young adult faces, we showed participants 40 undistorted faces from each of two categories and asked participants to rate each face on a 7-point attractiveness scale. In Experiment 2a, Caucasian and Chinese participants rated the attractiveness of Caucasian and East Asian faces; in Experiment 2b, young and older adults rated the attractiveness of young and older adult faces. To quantify consensus, for each face, we calculated the standard deviation of ratings across participants. We hypothesized that there would be less between-rater variability (i.e., smaller mean standard deviation) in ratings of individual faces for own-race faces (Experiment 2a) and young adult faces (Experiment 2b) relative to other-race and older adult faces, respectively.

### Experiment 2a

#### Method

**Participants.** Forty Caucasian undergraduates (37 female; mean age = 20.48 years,  $SD = 5.67$ , age range = 17–25) from Brock University, Canada and 40 Chinese



undergraduates (35 female; mean age = 22.15 years,  $SD = 2.43$ , age range = 18–26) from Zhejiang Normal University, China participated in this experiment. As in Experiment 1, Caucasian and Chinese participants reported very little contact with other-race individuals. All participants included in our analyses reported having less than two other-race friends; 98.75% of participants reported having zero other-race friends. An additional 19 participants (two Chinese, 17 Caucasian) were excluded from the final analysis due to reported significant experience with individuals of other-race ethnicity. Participants received either research credit or a small honorarium for their participation.

**Materials.** Stimuli comprise colored photographs of 40 Caucasian faces (20 female) and 40 East Asian faces (20 female). All stimuli were acquired from the Center for Vital Longevity Face Database (Minear & Park, 2004) and from the Let's Face It database at Brock University. Each face was presented in a frontal view and with a neutral expression. Using Adobe Photoshop Version CS5, we removed the neck, background details, and distracting blemishes from the original pictures and resized them such that the distance from the hairline to the chin was approximately 500 pixels. All stimuli were presented and responses were recorded using SuperLab 4.5 software.

**Procedure.** This study received clearance from the Research Ethics Board at Brock University. After providing written informed consent, participants sat 60 cm in front of a 23-in. computer and were told that they would be shown a series of faces and that it was their job to rate each face in terms of its attractiveness. Participants were told that they would use a 7-point attractiveness rating scale, with 1 being *not at all attractive* and 7 being *extremely attractive*. Participants were told to attempt to use the full range of the scale and to think about the attractiveness of each face with regard to other faces of that race when making their responses.

Face race was blocked such that half of the participants viewed Caucasian faces followed by East Asian faces, and half viewed East Asian faces followed by Caucasian faces. Each block contained 40 trials and each trial consisted of a 500-ms fixation cross followed by a face that appeared for 3,000 ms. Participants had an unlimited amount of time to rate each face's attractiveness via keypad on the 7-point scale. Before each block, participants were presented with all 40 faces from that block, one at a time for 1 s each, with a 500-ms ISI. This was done so that participants would have a sense of the range of variability in the attractiveness of the faces, thus ensuring that the first few faces would not be given abnormal ratings. As in Experiment 1, upon completion of the attractiveness task, participants completed a questionnaire assessing the amount of contact they had with other-race identities.

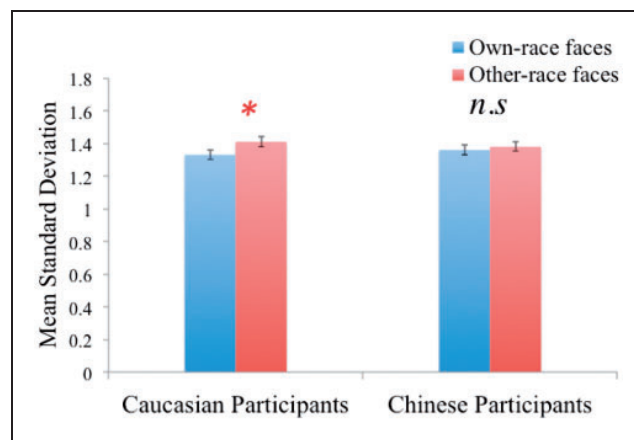
**Results and discussion.** For each Caucasian and East Asian face, we calculated the mean attractiveness and the standard deviation in attractiveness ratings; calculations were done separately for Caucasian and Chinese participants. The mean attractiveness rating for each face reflects the average (i.e., central tendency) rating provided by Caucasian or Chinese raters. The standard deviation in attractiveness ratings reflects the extent to which raters agree with each other regarding the attractiveness of a particular face. In other words, higher standard deviations in attractiveness ratings indicate greater between-rater variability (i.e., less consensus) in attractiveness ratings.

Task order and face sex did not have a significant effect on the mean and  $SD$  attractiveness ratings, nor did they interact with any other variables (all  $ps > .09$ ); thus, task order and face sex were excluded in all subsequent analyses. All follow-up  $t$  tests were two-tailed.

**Mean attractiveness ratings.** A 2 (face race: Caucasian, East Asian)  $\times$  2 (participant race: Caucasian, Chinese) **mixedmodel** ANOVA examining mean attractiveness ratings for own- and other-race faces revealed no main effects of face race,  $F(1, 78) = 1.12, p = .29, \eta_p^2 = .01$ , and participant race,  $F(1, 78) = 0.32, p = .58, \eta_p^2 = .004$ . Furthermore, there was no significant face race by participant race interaction,  $F(1, 78) = 2.23, p = .14, \eta_p^2 = .03$ , indicating that Caucasian and Chinese adults showed comparable average attractiveness ratings for both own- and other-race faces. This was further confirmed by a significant positive correlation between Caucasian and Chinese participants' mean attractiveness ratings for both Caucasian ( $r = 0.61, p < .001$ ) and East Asian faces ( $r = 0.62, p < .001$ ). In other words, increases in Caucasian raters' mean attractiveness rating for Caucasian and East Asian faces were associated with increases in Chinese raters' mean attractiveness rating for the same faces.

**Standard deviation in attractiveness ratings.** We next examined whether Caucasian and Chinese raters showed greater between-participant variability (e.g., less consensus) in rating the attractiveness of other-race relative to own-race faces. We conducted a 2 (face race: Caucasian, East Asian)  $\times$  2 (participant race: Caucasian, Chinese) **mixedmodel** ANOVA with the standard deviations in attractiveness ratings for own- and other-race faces as the dependent variable. We found a significant interaction between face race and participant race,  $F(1, 78) = 5.24, p = .03, \eta_p^2 = .06$  (see Figure 3). Independent-sample  $t$  tests confirmed that Caucasian adults showed greater between-participant variability when rating the attractiveness of other-race faces ( $M = 1.41, SE = 0.03$ ) than when rating the attractiveness of own-race faces ( $M = 1.33, SE = 0.03; t_{78} = 2.18, p = .03, \text{Cohen's } d = 0.47$ ). However, Chinese adults showed comparable between-participant variability when rating the attractiveness of both own- ( $M = 1.36, SE = 0.03$ ) and other-race faces ( $M = 1.39, SE = 0.03; t_{78} = -0.64, p = .53, \text{Cohen's } d = 0.16$ ).

In summary, consistent with our hypothesis, Caucasian participants showed reduced consensus when rating the attractiveness of East Asian faces compared to Caucasian faces. Previous studies suggest that perceived facial attractiveness reflects norm-based coding, with attractiveness ratings inversely related to the distance from the norm (e.g., Rhodes & Tremewan, 1996). Among Caucasian participants, reduced consensus when judging the attractiveness of other-race faces is consistent with their impaired ability to judge the



Please use the updated figure here

**Figure 3.** Caucasian and Chinese adults' mean standard deviation in attractiveness ratings for own- and other-race faces. \* $p < .05$ , n.s. = nonsignificant.

normality of other-race faces (Experiment 1). In contrast, despite showing higher accuracy when judging the normality of East Asian compared to Caucasian faces (Experiment 1), Chinese participants did not show reduced consensus when judging the attractiveness of Caucasian faces. This might be attributable to Chinese participants having greater exposure to Western media (e.g., Western movies, singers, news) than Caucasian participants have to East Asian media. To the extent that actors and singers are typically above average in attractiveness, this experience might influence Chinese participants' judgments of attractiveness more so than their judgments of normality, a possibility that should be examined in future studies. Collectively, the results of Experiments 1 and 2a provide direct evidence that the dimensions of face space are more refined for own- than other-race faces. In Experiment 2b, we wanted to confirm these findings with another face category with which adults have differential experience: young and older adult faces. In particular, we examined whether consensus in attractiveness ratings varies as a function of face age.

### *Experiment 2b*

Short and Mondloch (2013) reported that both young and older adults are more accurate in judging the normality of young relative to older adult faces under conditions in which their ability to discriminate faces from the two categories was comparable. In Experiment 2b, we measured the degree of consensus among young and older adults when judging the attractiveness of young versus older adult faces. Greater consensus when judging young adult faces would provide converging evidence of a more refined norm for young faces, consistent with the dimensions of face space being optimized for young adult faces.

#### *Method*

**Participants.** Forty Caucasian undergraduate students from Brock University (35 female;  $M = 19.60$  years, age range = 18–24) and 40 senior citizens living in independent housing in the Niagara region of Ontario (29 female;  $M = 71.88$  years, age range = 60–89) participated in this experiment. Senior citizen participants were all in good health, and 39 of the 40 senior participants had 20/30 vision or better. Undergraduates received research credit or a small honorarium and senior citizens received a gift card for their participation in the study. All participants completed a questionnaire assessing their weekly face-to-face contact with both young and older adults. All participants included in our analyses reported spending more time with own-age peers ( $M = 45.24$  hours and 53.73 hours per week for young and older adults, respectively) than other-age individuals ( $M = 7.63$  hours and 7.69 hours per week). An additional two undergraduates were tested but excluded from the final data set because they failed to pay attention during testing ( $n = 1$ ) or did not fill out the questionnaire properly ( $n = 1$ ).

**Materials.** Stimuli comprise colored photographs of 40 Caucasian young adult (20 female; age range = 18–29) and 40 Caucasian older adult (20 female; age range = 70–81) faces. All stimuli were acquired from the Center for Vital Longevity Face Database (Minear & Park, 2004) and resized such that the distance from the hairline to the chin was approximately 500 pixels. Young adult stimuli were identical to those used in Experiment 2a. As in Experiment 2a, all photographs were cropped such that only the face and hair remained and all distracting blemishes were removed. All stimuli were presented and responses were recorded using SuperLab 4.5 software.

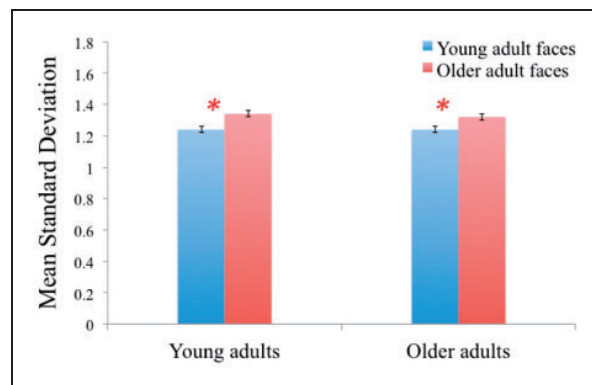
**Procedure.** The procedure of Experiment 2b was identical to that of Experiment 2a but older adult faces were shown instead of East Asian faces and the questionnaire measured the amount of current contact with young versus older adult faces.

**Results and discussion.** Similar to Experiment 2a, we calculated the mean attractiveness rating and standard deviation for each young and older adult face.

**Mean attractiveness ratings.** A 2 (face age: young, older)  $\times$  2 (participant age: young, older) **mixedmodel** ANOVA with mean attractiveness ratings as the dependent variable revealed a main effect of participant age,  $F(1, 78) = 86.76, p < .001, \eta_p^2 = .53$ , such that faces were rated as more attractive by older adults ( $M = 4.32, SE = .07$ ) than by young adults ( $M = 3.73, SE = .10$ ). There was no main effect of face age,  $F(1, 78) = 2.34, p = .13, \eta_p^2 = .03$ , nor was there a significant face age by participant age interaction,  $F(1, 78) = .17, p = .68, \eta_p^2 = .002$ . The lack of a significant interaction indicates that young and older adults provided comparable attractiveness ratings for both face ages; this was further confirmed by a significant positive correlation between young and older adults' mean attractiveness ratings for young adult ( $r = .77, p < .001$ ) and older adult faces ( $r = .77, p < .001$ ).

**Standard deviation in attractiveness ratings.** A 2 (face age: young, older)  $\times$  2 (participant age: young, older) **mixedmodel** ANOVA with the standard deviations in attractiveness ratings as the dependent variable revealed a main effect of face age,  $F(1, 78) = 13.96, p < .001, \eta_p^2 = .15$  (see Figure 4). Overall, there was greater between-participant variability in attractiveness ratings for older adult faces ( $M = 1.32, SE = .02$ ) than for young adult faces ( $M = 1.24, SE = .02$ ). There was no main effect of participant age,  $F(1, 78) = .24, p = .63, \eta_p^2 = .003$ , nor a significant face age by participant age interaction,  $F(1, 78) = .55, p = .46, \eta_p^2 = .007$ .

In summary, just as Caucasian adults showed greater consensus in their attractiveness ratings of own- relative to other-race faces, both young and older adults showed greater consensus in their attractiveness ratings for young relative to older adult faces. This is consistent with previous evidence that young and older adults are more sensitive to deviations from the norm in young relative to older faces (Short & Mondloch, 2013). It is surprising, perhaps, that older adults showed an advantage for young adult faces in this task despite recent abundant experience with older adult faces. This is in contrast to Chinese young adults who, likely because of exposure to Western media, did not show an own-race advantage on our attractiveness task. We suspect that the continuous young adult face bias in



Please use the updated figure here

**Figure 4.** Young and older adults' mean standard deviation in attractiveness ratings for young and older adult faces. \* $p < .05$ .

older adults reflects the special influence of early experience in shaping perceptual expertise—experience that is dominated by young faces (Macchi Cassia, Bulf, Quadrelli, & Proietti, 2013; Short, Semplonius, Proietti, & Mondloch, 2014). Collectively, these results provide direct evidence of a more refined face space (a well-defined norm and sensitivity to deviations along the underlying dimensions) for own-race and young adult faces relative to faces from other categories.

## General Discussion

Collectively, our results provide the first direct evidence that multidimensional face space is more refined for own- than other-race faces and provide converging evidence that both young and older adults' face space is more refined for young than older adult faces (see Short & Mondloch, 2013). We discovered an own-race advantage in judgments of normality but not discrimination (Experiment 1) and among Caucasian participants, less between-participant variability in attractiveness ratings for own-race than for other-race faces (Experiment 2a). We also found less between-participant variability in attractiveness ratings for young than for older adult faces both among young and older participants (Experiment 2b). These results suggest that the dimensions of face space are optimized for face categories with which people have ample perceptual experience (i.e., own-race faces; young adult faces). A particular strength in this set of experiments is that we tested adults from different race and age groups; consequently, variability in normality and attractiveness judgments cannot be attributed to stimulus effects.

It has been well established that other-race faces and older adult faces are recognized less accurately than own-race and young adult faces (Bothwell et al., 1989; de Heering & Rossion, 2008; Meissner & Brigham, 2001). This has been attributed to norm-based coding, but few studies have systematically examined the relationship between norm-based coding and recognition deficits. Although opposing aftereffects suggest that separable norms are used to encode faces from different categories, the hypothesis that the face norm and the dimensions underlying face space are less well differentiated for faces from less encountered categories has not been directly tested in studies of aftereffects. Here, we provide the first direct evidence of this hypothesis.

Unlike judging which of two faces is more expanded, judging normality requires perceptual expertise (i.e., knowledge of what an average face from that category looks like). This argument is supported by evidence that inversion impairs performance on the normality task (but not the discrimination task) and eliminates the young adult face advantage (Short & Mondloch, 2013). In the current study, despite no difference in the accuracy with which Caucasian and Chinese adults were able to discriminate own- versus other-race faces, their judgments of normality were more accurate for own-race faces. This result reflects reduced sensitivity to deviations from a prototypical other-race face and inefficiency in the use of norm-based coding for other-race faces, a pattern likely resulting from limited perceptual experience with other-race faces. This finding is consistent with evidence that young and older adults are more sensitive to deviations from normality in young than older adult faces (Short & Mondloch, 2013), suggesting a reliance on a face space that is optimized for the dimensions for young adult faces.

In Experiment 2, we provided converging evidence for less efficiently tuned dimensions of face space for other-race and older adult faces by using a different task: attractiveness judgments. We found that Caucasian adults showed more consensus (less variability) when rating the attractiveness of Caucasian faces compared to East Asian faces, with no difference among Chinese adults. Moreover, both young and older adults were more likely to agree on

the attractiveness of young adult faces than older adult faces. Greater between-participant variability in perceived attractiveness for other-race and older adult faces is consistent with our conclusion that the perceptual processing system is preferentially tuned for face categories with which people have more perceptual experience.

The results of the present study support the norm-based coding model of the own-race and own-age/young adult face recognition advantage and highlight the important role of perceptual experience in shaping the face norm and the dimensions underlying face space. The dimensions of face space are refined through perceptual experience to represent the facial properties that are optimal for discriminating identities from highly familiar categories. Consequently, faces from unfamiliar categories, such as other-race faces and older adult faces, are tightly clustered in the periphery of face space (Valentine, 1991). This model explains why perceivers have an impaired ability to detect deviations from normality and greater between-perceiver variability in attractiveness ratings for other-race and older adult faces. This may be one reason why participants make more errors in recognition tasks involving other-race and older adult identities (e.g., Golby, Gabrieli, Chiao & Eberhardt, 2001; MacLin & Malpass, 2001; Wright, Boyd & Tredoux, 2003).

Our finding that there is a perceptual advantage for own-race and young adult faces is consistent with evidence that both N170 amplitude and the N170 inversion effect are influenced by face race and age. The amplitude of the N170 is smaller for own- than other-race faces (e.g., Wiese, Kaufmann, & Schweinberger, 2014) and for young than older faces in both young and older adults (Wiese, Schweinberger, & Hansen, 2008), whereas the N170 inversion effect shows the opposite pattern (Komes, Schweinberger, & Wiese, 2015; Vizioli, Rousselet, & Caldara, 2010; Wiese, Komes, & Schweinberger, 2013). This early perceptual advantage impacts recognition; the own-race recognition advantage is robust, as is the own-age advantage among young adults. In contrast, findings are inconclusive in older adult samples (see Proietti, Macchi Cassia, & Mondloch, 2015 for a review), perhaps because of later processing stages (e.g., as reflected in the N250; Wiese, Kachel, & Schweinberger, 2013) being influenced by accumulation of experience with different age groups over the lifespan (Anastasi & Rhodes, 2006) and the special influence of early experience in shaping perceptual expertise for adult faces (Macchi Cassia et al., 2013).

Our finding that face space is more refined for own-race and young adult faces also has explanatory value for a less investigated challenge in face recognition: recognizing identity when appearance varies. Two pictures of the same person can look very different and pictures of two different people can look very similar. When sorting photographs of unfamiliar faces into piles such that each pile includes all of the pictures of one identity, adults frequently separate photos of one person into multiple piles (i.e., they perceive different pictures of the same person as belonging to different identities). For example, when sorting a pile of 40 photographs **comprises 20 pictures** of two different identities, adults make about seven piles (i.e., they perceive about seven different identities; Jenkins, White, Van Montfort, & Burton, 2011). They make twice as many piles (perceive twice as many identities) when sorting unfamiliar other-race identities (Laurence, Zhou, & Mondloch, 2015), suggesting that recognizing identity in photos that capture natural within-person variability in appearance among other-race faces is especially challenging.

This finding was interpreted in light of extensions of Valentine's norm-based coding model, according to which each face is represented as a region (attractor field), rather than a single point, in face space (Tanaka, Giles, Kremen, & Simon, 1998). The attractor field reflects the range of inputs that are perceived as belonging to a given identity (i.e., our ability to tolerate within-person variability in appearance). The size of an identity's attractor field is inversely correlated with the density of nearby representations, and thus hypothesized to be smaller for

other-race and older adult faces than for own-race and young adult faces. Laurence et al. (2015) argued that smaller attractor fields for other-race faces not only make other-race faces harder to tell apart but increase the difficulty in recognizing an identity in the context of natural changes in appearance. The current study provides evidence that other-race faces are more densely clustered than own-race faces in face space, with a smaller inter-face distance—a key component to this argument. It would be worthwhile to investigate the relationship between individual differences in sensitivity to deviations from the norm and the ability to recognize pictures of faces that incorporate a wide range of natural variations.

### *Issues for Future Research*

Our findings provide direct evidence that deficits in recognizing other-race and older adult faces can be attributed to a less refined face space for faces from these categories. They also raise several issues worthy of further investigation. In particular, we highlight the need to refine our conceptualization of face space. First, it is not clear exactly what the dimensions underlying face space are. They might be features and their spacing (e.g., nose length, distance between the eyes) or more abstract dimensions (e.g., eigenfaces; Hancock, Burton, & Bruce, 1996). Although norm-based coding has enormous explanatory power, it is important to better specify the nature of the underlying dimensions.

Second, the process through which perceptual experience shapes the dimensions of face space has not been specified. Opposing aftereffects suggest that we have separable face spaces for own- and other-race faces (Jaquet et al., 2008) and young versus older faces (Short et al., 2015). Partial transfer of aftereffects across face race (Jaquet & Rhodes, 2008) and age (Short et al., 2015) suggests shared underlying dimensions and separable prototypes/norms. Some dimensions are almost certainly shared across categories (Short et al., 2015). This characterization of face space accounts for our findings in several ways. First, it is likely that any one dimension is not equally diagnostic for faces from all categories. To the extent that dimensions are optimized to discriminate faces from categories with which people have more perceptual experience (e.g., own-race and young adult faces), they will be less effective for discriminating faces from other categories (e.g., other-race and older adult faces). For example, eye color may be a salient dimension for discriminating Caucasian identities, but Asian identities might be more clustered on this dimension. In addition, we propose that regions of face space associated with different categories vary in the number of dimensions represented and/or the length of the underlying vectors (a conceptualization of sensitivity to differences along dimensions). Just as children rely on fewer dimensions than adults (Nishimura, Maurer, & Gao, 2009), it is likely that the very limited number of other-race and older adult face exemplars in one's face space severely restricts the number of underlying dimensions. As noted by Burton and Vokey (1998), fewer dimensions might leave the vast majority of faces clustered in the center of face space. As dimensions are added (which happens throughout development for own-race, young adult faces), faces become more dispersed, making them easier to discriminate and recognize. Our results suggest that this dispersion also leads to greater consensus in attractiveness judgments because there is more variability among faces in their proximity to the center of face space. While in contrast to Valentine's (1991) claim that faces are most densely clustered in the center of face space, it is consistent with his argument that perceived attractiveness is influenced by distance from the center. Future studies should aim to clarify how representations vary across face categories.

There is also evidence showing that children are less sensitive than adults in differentiating along the dimensions of face space (Anzures, Mondloch, & Lackner, 2009; Short, Hatry, & Mondloch, 2011; Short, Lee, Fu, & Mondloch, 2014) and their ability to simultaneously use

multiple dimensions improves after 8 years of age (Nishimura et al., 2009). Future studies should investigate whether the increase of perceptual expertise with age enhances children's sensitivity to deviations from the face norm and their sensitivity to the dimensions along which faces vary.

In summary, two methodologies were used to examine the representation of own- versus other-race faces and young versus older adult faces in face space. Adults were more sensitive to how faces deviate from an average face when judging own- relative to other-race faces and were less likely to agree on the attractiveness of other-race and older adult faces. Collectively, these results provide direct evidence that perceptual experience with own-race, young adult faces optimizes the dimensions of face space for own-race and young adult faces. Such reduced sensitivity to deviations from the face norms for categories with which we have less experience may explain the special challenge of recognizing other-race and older adult identities.

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

### References

- Alley, T. R., & Cunningham, M. R. (1991). Averaged faces are attractive, but very attractive faces are not average. *Psychological Science*, 2, 123–125. doi:10.1111/j.1467-9280.1991.tb00113.x
- Anastasi, J. S., & Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin & Review*, 12, 1043–1047. doi:10.3758/BF03206441
- Anastasi, J. S., & Rhodes, M. G. (2006). Evidence for an own-age bias in face recognition. *North American Journal of Psychology*, 8, 237–253.
- Anzures, G., Mondloch, C. J., & Lackner, C. (2009). Face adaptation and attractiveness aftereffects in 8-year-olds and adults. *Child Development*, 80, 178–191. doi:10.1111/j.1467-8624.2008.01253.x
- Bothwell, R. K., Brigham, J. C., & Malpass, R. S. (1989). Cross-racial identification. *Personality and Social Psychology Bulletin*, 15, 19–25. doi:10.1177/0146167289151002
- Burton, A. M., & Vokey, J. R. (1998). The face-space typicality paradox: Understanding the face-space metaphor. *The Quarterly Journal of Experimental Psychology: Section A*, 51, 475–483. doi:10.1080/713755768
- Byatt, G., & Rhodes, G. (1998). Recognition of own-race and other-race caricatures: Implications for models of face recognition. *Vision Research*, 38, 2455–2468. doi:10.1016/S0042-6989(97)00469-0
- Cross, J. F., & Cross, J. (1971). Age, sex, race, and the perception of facial beauty. *Developmental Psychology*, 5, 433–439. doi:10.1037/h0031591
- Cunningham, M. R., Roberts, A. R., Barbee, A. P., Druen, P. B., & Wu, C. H. (1995). “Their ideas of beauty are, on the whole, the same as ours”: Consistency and variability in the cross-cultural perception of female physical attractiveness. *Journal of Personality and Social Psychology*, 68, 261. doi:10.1037/0022-3514.68.2.261
- de Heering, A., & Rossion, B. (2008). Prolonged visual experience in adulthood modulates holistic face perception. *PLoS ONE*, 3, e2317. doi:10.1371/journal.pone.0002317
- Dennett, H. W., McKone, E., Edwards, M., & Susilo, T. (2012). Face aftereffects predict individual differences in face recognition ability. *Psychological Science*, 23, 1279–1287. doi:10.1177/0956797612446350



- Golby, A. J., Gabrieli, J. D., Chiao, J. Y., & Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. *Nature Neuroscience*, *4*, 845–850. doi:10.1038/90565
- Hancock, P. J., Burton, A. M., & Bruce, V. (1996). Face processing: Human perception and principal components analysis. *Memory & Cognition*, *24*, 26–40. doi:10.3758/BF03197270
- Jaquet, E., & Rhodes, G. (2008). Face aftereffects indicate dissociable, but not distinct, coding of male and female faces. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 101–112. doi:10.1037/0096-1523.34.1.101
- Jaquet, E., Rhodes, G., & Hayward, W. G. (2008). Race-contingent aftereffects suggest distinct perceptual norms for different race faces. *Visual Cognition*, *16*, 734–753. doi:10.1080/12506280701350647
- Jenkins, R., White, D., Van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. *Cognition*, *121*, 313–323. doi:10.1016/j.cognition.2011.08.001
- Komes, J., Schweinberger, S. R., & Wiese, H. (2015). Neural correlates of cognitive aging during the perception of facial age: The role of relatively distant and local texture information. *Frontiers in psychology*, *6*, 1–16. doi:10.3389/fpsyg.2015.01420
- Kuefner, D., Macchi Cassia, V., Picozzi, M., & Bricolo, E. (2008). Do all kids look alike? Evidence for an other-age effect in adults. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 811–817. doi:10.1037/0096-1523.34.4.811
- Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological bulletin*, *126*, 390–423. doi:10.1037//0033-2909.126.3.390
- Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science*, *1*, 115–121. doi:10.1111/j.1467-9280.1990.tb00079.x
- Laurence, S., Zhou, X., & Mondloch, C. J. (2015). The flip side of the other-race coin: They all look different to me. *British Journal of Psychology*, *107*, 374–388. doi:10.1111/bjop.12147
- Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience*, *4*, 89–94. doi:10.1038/ncb739
- Light, L. L., Hollander, S., & Kayra-Stuart, F. (1981). Why attractive people are harder to remember. *Personality and Social Psychology Bulletin*, *7*, 269–276. doi:10.1177/014616728172014
- Little, A. C., DeBruine, L. M., & Jones, B. C. (2005). Sex-contingent face after-effects suggest distinct neural populations code male and female faces. *Proceedings of the Royal Society of London B: Biological Sciences*, *272*, 2283–2287. doi:10.1098/rspb.2005.3220
- Little, A. C., DeBruine, L. M., Jones, B. C., & Waitt, C. (2008). Category contingent aftereffects for faces of different races, ages and species. *Cognition*, *106*, 1537–1547. doi:10.1016/j.cognition.2007.06.008
- Macchi Cassia, V., Bulf, H., Quadrelli, E., & Proietti, V. (2013). Age-related face processing bias in infancy: Evidence of perceptual narrowing for adult faces. *Developmental psychobiology*, *56*, 238–248. doi:10.1002/dev.21191
- MacLin, O. H., & Malpass, R. S. (2001). Racial categorization of faces: The ambiguous race face effect. *Psychology, Public Policy, and Law*, *7*, 98. doi:10.1037//1076-8971.7.1.98
- Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, *7*, 3. doi:10.1037/1076-8971.7.1.3
- Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*, *36*, 630–633. doi:10.3758/BF03206543
- Morris, P. E., & Wickham, L. H. (2001). Typicality and face recognition: A critical re-evaluation of the two factor theory. *The Quarterly Journal of Experimental Psychology: Section A*, *54*, 863–877. doi:10.1080/713755992
- Nishimura, M., Maurer, D., & Gao, X. (2009). Exploring children's face-space: A multidimensional scaling analysis of the mental representation of facial identity. *Journal of Experimental Child Psychology*, *103*, 355–375. doi:10.1016/j.jecp.2009.02.005
- 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*, 208–224. doi:10.3758/BF03208892

- Perfect, T. J., & Harris, L. J. (2003). Adult age differences in unconscious transference: Source confusion or identity blending? *Memory & cognition*, *31*, 570–580. doi:10.3758/BF03196098
- Proietti, V., Macchi Cassia, V., & Mondloch, C. J. (2015). The own-age face recognition bias is task dependent. *British Journal of Psychology*, *106*, 446–467. doi:10.1111/bjop.12104
- Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology*, *57*, 199–226. doi:10.1146/annurev.psych.57.102904.190208
- Rhodes, M. G., & Anastasi, J. S. (2012). The own-age bias in face recognition: A meta-analytic and theoretical review. *Psychological Bulletin*, *138*, 146–174. doi:10.1037/a0025750
- Rhodes, G., & Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision Research*, *46*, 2977–2987. doi:10.1016/j.visres.2006.03.002
- Rhodes, G., Jeffery, L., Taylor, L., Hayward, W. G., & Ewing, L. (2014). Individual differences in adaptive coding of face identity are linked to individual differences in face recognition ability. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 897. doi:10.1037/a0035939
- 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*, 558–566. doi:10.1046/j.0956-7976.2003.psci\_1465.x
- Rhodes, G., Jeffery, L., Watson, T. L., Jaquet, E., Winkler, C., & Clifford, C. W. G. (2004). Orientation-contingent face aftereffects and implications for face-coding mechanisms. *Current Biology*, *14*, 2119–2123. doi:10.1016/j.cub.2004.11.053
- Rhodes, G., & Leopold, D. A. (2011). Adaptive norm-based coding of face identity. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 263–286). Oxford, England: Oxford University Press.
- Rhodes, G., Robbins, R., Jaquet, E., McKone, E., Jeffery, L., & Clifford, C. W. G. (2005). Adaptation and face perception: How aftereffects implicate norm-based coding of faces. *Fitting the mind to the world: Adaptation and after-effects in high-level vision* [AQ2] (pp. 213–240). New York: Oxford University Press.
- Rhodes, G., & Tremewan, T. (1996). Averageness, exaggeration, and facial attractiveness. *Psychological Science*, *7*, 105–110. doi:10.1111/j.1467-9280.1996.tb00338.x
- Rhodes, G., Watson, T. L., Jeffery, L., & Clifford, C. W. (2010). Perceptual adaptation helps us identify faces. *Vision Research*, *50*, 963–968. doi:10.1016/j.visres.2010.03.003
- Schweinberger, S. R., Zäske, R., Walther, C., Golle, J., Kovács, G., & Wiese, H. (2010). Young without plastic surgery: Perceptual adaptation to the age of female and male faces. *Vision Research*, *50*, 2570–2576. doi:10.1016/j.visres.2010.08.017
- Short, L. A., Hatry, A. J., & Mondloch, C. J. (2011). The development of norm-based coding and race-specific face prototypes: An examination of 5- and 8-year-olds' face space. *Journal of Experimental Child Psychology*, *108*, 338–357. doi:10.1016/j.jecp.2010.07.007
- Short, L. A., Lee, K., Fu, G., & Mondloch, C. J. (2014). Category-specific face prototypes are emerging, but not yet mature, in 5-year-old children. *Journal of Experimental Child Psychology*, *126*, 161–177. doi:10.1016/j.jecp.2014.04.004
- Short, L. A., & Mondloch, C. J. (2013). Aging faces and aging perceivers: Young and older adults are less sensitive to deviations from normality in older than in young adult faces. *Perception*, *42*, 795–812. doi:10.1068/p7380
- Short, L. A., Mondloch, C. J., & Hackland, A. T. (2015). Attractiveness judgments and discrimination of mommies and grandmas: Perceptual tuning for young adult faces. *Journal of Experimental Child Psychology*, *129*, 1–11. doi:10.1016/j.jecp.2014.08.001
- Short, L. A., Proietti, V., & Mondloch, C. J. (2015). Representing young and older adult faces: Shared or age-specific prototypes? *Visual Cognition*, *23*, 939–956. doi:10.1080/13506285.2015.1115794
- Short, L. A., Semplonius, T., Proietti, V., & Mondloch, C. J. (2014). Differential attentional allocation and subsequent recognition for young and older adult faces. *Visual Cognition*, *22*, 1272–1295. doi:10.1080/13506285.2014.993007
- Tanaka, J., Giles, M., Kremen, S., & Simon, V. (1998). Mapping attractor fields in face space: The atypicality bias in face recognition. *Cognition*, *68*, 199–220. doi:10.1016/S0010-0277(98)00048-1

- Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology A*, *43*, 161–204. doi:10.1080/14640749108400966
- Vizioli, L., Rousselet, G. A., & Caldara, R. (2010). Neural repetition suppression to identity is abolished by other-race faces. *Proceedings of the National Academy of Sciences*, *107*, 20081–20086. doi:10.1073/pnas.1005751107
- Webster, M. A., & MacLeod, D. I. A. (2011). Visual adaptation and face perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*, 1702–1725. doi:10.1098/rstb.2010.0360
- Wiese, H., Kachel, U., & Schweinberger, S. R. (2013). Holistic face processing of own-and other-age faces in young and older adults: ERP evidence from the composite face task. *Neuroimage*, *74*, 306–317. doi:10.1016/j.neuroimage.2013.02.051
- Wiese, H., Kaufmann, J. M., & Schweinberger, S. R. (2014). The neural signature of the own-race bias: Evidence from event-related potentials. *Cerebral Cortex*, *24*, 826–835. doi:10.1093/cercor/bhs369
- Wiese, H., Komes, J., & Schweinberger, S. R. (2013). Ageing faces in ageing minds: A review on the own-age bias in face recognition. *Visual Cognition*, *21*, 1337–1363. doi:10.1080/13506285.2013.823139
- Wiese, H., Schweinberger, S. R., & Hansen, K. (2008). The age of the beholder: ERP evidence of an own-age bias in face memory. *Neuropsychologia*, *46*, 2973–2985. doi:10.1016/j.neuropsychologia.2008.06.007
- Wright, D. B., Boyd, C. E., & Tredoux, C. G. (2003). Inter-racial contact and the own-race bias for face recognition in South Africa and England. *Applied Cognitive Psychology*, *17*, 365–373. doi:10.1002/acp.898.