Impact of Bi-ethnic Exposure on Face Memory and Perception

by

Seyed Morteza Mousavi

B.Sc., The University of British Columbia 2017

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL

STUDIES

(Neuroscience)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2019

©Seyed Morteza Mousavi, 2019

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, a thesis/dissertation entitled:

Impact of Bi-ethnic Exposure on Face Memory and Perception

submitted by Seyed Morteza Mousavi in partial fulfillment of the requirements for

the degree of Master of Science

in Neuroscience

Examining Committee:

Dr. Ipek Oruc, Ophthalmology and Visual Sciences

Supervisor

Dr. James Enns, Psychology

Supervisory Committee Member

Dr. Janet Werker, Psychology

Supervisory Committee Member

Dr. Randi Starrfelt, Psychology, University of Copenhagen

Additional Examiner

Abstract

Tuning changes of perception from a generalist to a specialist system is termed perceptual narrowing (Werker & Tees, 1984). Observers raised in ethnically homogenous environments show difficulties in recognizing other-race faces (Meissner & Brigham, 2001). Termed the other-race effect (ORE), this phenomenon is in part attributed to lack of visual experience with other-race faces (Chiroro & Valentine, 1995). The ORE is a form of perceptual narrowing that allows specialization for commonly encountered faces. To examine the origin of perceptual narrowing in face recognition we postulate three competing hypotheses. Based on the *Resource Bottleneck* hypothesis, limited neural and visual resources must be focused to achieve refined processing of own-race faces. Alternatively, the *Use/Disuse* hypothesis suggests that the ORE is merely due to lack of exposure necessary to reach expertise for unfamiliar face classes. Lastly, the *Facilitation* hypothesis, suggests that exposure to a diverse set of faces can lead to richer face representations, and thus, advantages in face recognition. To evaluate these hypotheses, we compared face abilities of 'Bi-ethnics'--individuals with sustained high exposure to both East Asian and Caucasian faces, to 'Mono-ethnics'—Caucasian and East Asian individuals. Face memory was examined using an old/new task for East Asian and Caucasian faces, as well as upright and inverted Cambridge Memory Tests for East Asian (McKone et al., 2012) and Caucasian (Duchaine & Nakayama, 2006) faces, in comparison to cars (Dennett et al., 2011) as a control stimulus. Lastly, an identity discrimination test of East Asian and Caucasian celebrities was used to examine face perception. Mono-ethnics showed ORE in memory tasks, whereas Bi-ethnics showed near-native level performance for both face categories. Mono-ethnics showed a larger face inversion effect for own-race faces and Bi-ethnics showed similar face inversion patterns for own-race in Mono-ethnics for both face

categories. All groups performed similarly in the Caucasian face perception task. However, in the East Asian face task, Caucasians showed ORE, with Bi-ethnics showing a small advantage over East Asians. Results across all tasks rule out the *Resource Bottleneck* hypothesis, and are consistent with the *Use/Disuse* hypothesis. Perception results suggest the possibility of *Facilitation* for Bi-ethnic exposure.

Lay Summary

Other-race faces are perceived with more difficulty than own-race faces. This phenomenon, termed as the other-race effect (ORE), happens due to narrowing of broad perceptual abilities to specialize in recognizing faces that one has more contact with, known as perceptual narrowing. To evaluate the origin of perceptual narrowing in face recognition we postulate three competing hypotheses. The *Resource Bottleneck* hypothesis states that perceptual narrowing happens to avoid division of neural and visual resources. The *Use/Disuse* hypothesis states that ORE happens solely due to lack of exposure to other-race faces. The *Facilitation* hypothesis predicts that exposure to more than one face-ethnicity can lead to advantages in face recognition. Our research revealed that individuals with sustained high exposure to Caucasian and East Asian faces show near-native face memory, ruling out the Resource Bottleneck hypothesis. These results are consistent with the *Use/Disuse* hypothesis and hint at the possibility of *Facilitation* regarding face perception.

Preface

All experiments are based on work conducted in UBC's Neuroscience of Vision and Action (NOVA) Laboratory (ICORD, Vancouver General Hospital), supervised by Dr. Ipek Oruc.

I was responsible for recruiting participants and running the experiments. Experiment 2 was done using tests previously developed by Duchaine & Nakamura (2006), Dennett et al. (2012), and McKone et al. (2012). I was responsible for the design and coding of Experiment 3 by modifying existing codes in the NOVA lab to adapt them to the present study with guidance and assistance of my supervisor Dr. Oruc. I was solely responsible for coding Experiment 1. Dr. Oruc and I worked on data processing and analysis. I wrote the thesis in full, with guidance and edits from my supervisor Dr. Oruc.

The UBC Behavioural Research Ethics board approved all procedures related to this work (title: Robust visual recognition of high-level form in human observers). Ethics board certificate: # H11-01527, Vancouver General Hospital approval: # V11-01527. This work has been funded by an NSERC Discovery grant (RGPIN 402654-11) and a grant from the Canada Foundation for Innovation John R. Evans Leaders Fund.

Table of Contents

List of Figures

Acknowledgments

First and foremost, thank you so much Dr. Oruc for all your support and help and for continuing to believe in me even in my not very bright days. I am very delighted that I have had you as my supervisor and have learned and continue to learn a lot from you. Working with you has truly been one of the few merits of my life.

I would also like to thank you my committee members Dr. Janet Werker and Dr. James Enns for providing very valuable insights on this project. Also, thank you Dr. Randi Starrfelt for accepting to be part of my committee as the external examiner.

Thank you to Todd Kamensek for all the laughter and stressful moments we passed together. I could have not wished for a better lab-mate.

And thank you to my wonderful family. To my mom, thank you for always being there for me. To my dad, thank you for always believing me. And to Melika, for being the loveliest younger sister.

To my parents

1 Introduction

Several lines of evidence from cognitive psychology (Bruce, Doyle, Dench, & Burton, 1991; Yin, 1969), computational vision (Turk & Pentland, 1991), neuropsychology (Behrmann, Winocur, & Moscovitch, 1992; Damasio, A.R, Tranel, D., Damasio, 1990), neurophysiology (Desimone, 2008; Perrett, Hietanen, Oram, & Benson, 1992), and neuroimaging (Kanwisher, McDermott, & Chun, 1997) suggest that mechanisms underlying face recognition are distinct from general-purpose mechanisms dealing with non-face visual stimuli. Nonetheless, visual experience plays a crucial role in modulating this ability. For example, more visual experience with own-race (Malpass & Kravitz, 1969), own-age (Anastasi & Rhodes, 2005; Lamont, Stewart-Williams, & Podd, 2005), and upright (Yin, 1969) faces leads to recognition biases towards these type of faces. On the other hand, contact with other-race (Sangrigoli, Pallier, a-M Argenti, Ventureyra, & de Schonen, 2005), other-age (Kuefner, Cassia, Picozzi, & Bricolo, 2008), and inverted (Laguesse, Dormal, Biervoye, Kuefner, & Rossion, 2012) faces can alleviate the biases towards them. In addition to these specific biases, limited experience with faces in general can impact overall face recognition abilities. This has been supported by results from a study showing poorer performance in face recognition tasks in small-town participants compared to participants who grew up in more densely populated areas (Balas & Saville, 2015). Moreover, it is believed that individuals with autism fail to develop perceptual expertise for faces due to a general lack of exposure to faces (Oruc, Shafai, & Iarocci, 2018; Schultz, 2005).

How differential experience with faces shape our face recognition system continues to be a topic that requires further evaluation. It has been shown that newborns face recognition abilities is very broad compared to adults (Mondloch et al., 1999), however eventually this broad sensitivity narrows down to form adult expertise in face recognition. In

this study, we examine the possible reasons behind this narrowing by focusing on exposure to more than one face-ethnicity.

1.1 The Other-Race Effect

Human adults' ability to recognize faces is remarkable considering that all faces share a similar configuration (Hancock & Rhodes, 2008). However, this expertise does not extend to all face categories. Individuals are more prone to making recognition errors when viewing faces from races other than their own. This phenomenon is known as the other-race effect (ORE), and it is one of the most robust findings in the face recognition literature (Meissner & Brigham, 2001; Sporer, 2001). The ORE has been replicated across different studies (Bothwell, Brigham, & Malpass, 2007) and observed across participants from different racial groups (Meissner & Brigham, 2001). It is also evident from research that the ORE is not occurring due to the physiognomy of faces from a specific race (O'toole, Deffenbacher, Valentin, & Abdi, 1994).

1.1.1 The ORE and the Role of Experience

The ORE is explained, in part, by the contact hypothesis. This hypothesis proposes that the ORE happens due to limited contact with exemplars of other-race faces (Chiroro & Valentine, 1995; Furl, Phillips, & O'Toole, 2002). Empirical studies of everyday exposure to faces show that, indeed, own-race faces dominate the face diet of infants (Sugden, Mohamed-Ali, & Moulson, 2014) as well adults (Oruc, Shafai, Murthy, Lages, & Ton, 2018). In the context of the ORE, the contact hypothesis predicts that an increase in interaction with another race will lead to increased exposure to, and improved recognition abilities for, faces of that race. This has been supported by results from several studies that had shown training with other-race faces can lead to an immediate improvement in their recognition (Elliott, Wills, & Goldstein, 1973; Goldstein & Chance, 1985; Tanaka, Heptonstall, & Hagen,

2013). In addition to these results pertaining to short-term effects, more long-term impacts have been recorded as well. For example, in a study focusing on African American and Caucasian children from segregated and integrated schools, it was found that those in segregated schools showed a greater ORE (Feinman & Entwisle, 2006). Similarly, in an examination of ethnically Korean adults who had been adopted by Caucasian families between the ages 3 to 9, Sangrigoli et al. (2005) found that these individuals showed a pattern of performance similar to Caucasian and opposite to Korean controls, where they performed less accurately for Korean faces than Caucasian faces.

1.1.2 The ORE and Face Memory

The first empirical accounts of the ORE showed that memory for other-race faces is poorer than for own-race. In a 1969 study, Malpass and Kravitz examined Caucasian and African American participants' ability to recall previously learned faces when viewed among novel face identities and found that both groups have a better recall ability for own-race faces than other-race. Since then several other studies have found similar results using the same old/new paradigm with different face ethnicities and testing on different racial groups, where participants were tasked to memorize a number of faces first, and then in a test phase they were presented with a mixture of novel and old faces and asked to select which faces they had seen previously (e.g., Valentine & Endo, 1992, Japanese and Caucasian faces; Golby, Gabrieli, Chiao, & Eberhardt, 2001, Caucasian and African American faces; Wright, Boyd, & Tredoux, 2003, Caucasian and African faces; Tanaka & Pierce, 2009, Hispanic and African American faces).

The old/new paradigm has been criticized by some researchers as a method to assess face memory. Lindsay and Wells (1983) argued that this type of task does not reflect real-world situations and proposed an alternative method that asks participants to recall a

target identity from memory in a line-up of faces (Evans, Marcon, & Meissner, 2009; Jackiw, Arbuthnott, Pfeifer, Marcon, & Meissner, 2008). However, either approach do not exclusively measure memory for faces but rather memory for particular photographs. The development of the Cambridge Face Memory Test (CFMT; Duchaine & Nakayama, 2006) has helped tackle this issue. The CFMT has stimuli with variant view-point and lighting to specifically target for face memory abilities with high internal reliability. Furthermore, the development of the Cambridge Face Memory Test Chinese (CFMT Chinese; McKone et al., 2012) has provided researchers with a robust tool to evaluate the interaction between race and face memory.

1.1.3 The ORE and Perceptual Discrimination

The ORE is not limited to face memory. Walker and Tanaka (2003) used a same/different sequential matching task to evaluate face perception differences for own- and other-race faces. In each trial of their experiment participants were presented with a parent Caucasian or East Asian face followed by a mask and then the same face or an Asian-Caucasian morph and were asked to determine whether they had seen the same face before and after the mask. The results from the study showed that participants detected differences in own-race faces more accurately than other-race faces. Moreover, Walker and Hewstone (2006) used the same task as the previous study but with Caucasian and South Asian parent faces to evaluate the impact of race on perceptual abilities in Caucasian and South Asian participants in the UK. The results showed an ORE within Caucasian participants but no difference in discrimination performance between the two stimulus categories for South Asian participants.

1.2 The Face Inversion Effect

Inverted faces are more difficult to recognize than upright faces (Goldstein, 1964; Hochberg & Galper, 1967). Inversion impacts recognition of other mono-oriented object categories also, but impairment is greater for faces. The first evidence for this disproportionate effect came from a study conducted by Yin (1969). Yin compared recognition memory for upright and inverted faces and objects. In the upright condition faces were better recognized than other stimuli, however inverted faces were the most difficult stimuli to be recognized. The face inversion effect (FIE) has been reproduced widely by numerous other studies (for a review, see Valentine, 1988). The FIE is attributed to a type of face recognition known as configural coding where stimuli are recognized based on spatial relationships between their components. A common assumption is that upright faces are recognized based on information about their isolated physical features (component coding) as well as information about the spatial relationships of those features (configural coding) (Gilliain Rhodes, Hayward, & Winkler, 2006). In the inverted orientation, configural coding is disrupted and faces are recognized solely based on component coding (Valentine, 1988; Rossion & Gauthier,2002). Conversely, it has been shown that inversion does not impact the recognition of isolated features (Bartlett, Searcy, & Abdi, 2006) Furthermore, the FIE is, in part, believed to be a result of higher expertise with faces, since individuals must have some level of expertise to be able to extract and use the distinct relational information from a stimulus (Diamond & Carey, 1986).

1.2.1 The FIE in Recognition of Own and Other-Race Faces

Since more expertise in recognizing faces leads to an increase in the FIE for faces, one would expect that the differential experience with own-race and other-race faces would yield a similar effect. In this case, several studies have found a larger decrement for

recognition of inverted own-race faces than other-race faces. Rhodes et al. (1989) presented participants with 4 sets of either East Asian or Caucasian faces in a learning phase, and then in a test phase, they showed either upright or inverted faces to test their recognition memory and found a greater FIE for own-race faces. Moreover, Sangrigoli and de Schonen (2004a) found greater FIE for own- than other-race faces in 3- to 5-year-old Caucasian children. They presented the participants with upright and inverted faces of own-race and other-race and tested recognition abilities using a 2AFC task and by comparing the accuracies between the conditions found that the participants were more sensitive to inversion for own-race faces than other-race faces. To explore the effect of experience on the FIE, Hancock and Rhodes (2008) investigated the FIE in Caucasian and Chinese participants and found a decreased difference between own- and other-race FIE in those participants with higher cross-race contact.

In all, configural coding is a hallmark of expertise in visual recognition. We have more experience with faces than other mono-oriented object categories, hence we use this style of recognition more for faces. Within the realm of faces, greater contact with own-race faces, results in greater contribution of configural coding, leading to the empirical observation of greater FIE magnitude for own-race faces.

1.3 Perceptual Narrowing

Our tuning to sensory inputs is initially broad, enabling infants to process native and non-native sensory inputs similarly. During development, sensory tuning narrows as a function of selective native inputs leading to a decline in the integration of non-native stimuli (Lewkowicz & Ghazanfar, 2009). This phenomenon, termed *perceptual narrowing,* was first demonstrated in a study on consonant discrimination (Werker, Gilbert, Humphrey, & Tees, 1981). Building on earlier work by Eimas et al. (1971) that had shown infants have the

ability to discriminate non-native speech sounds, Werker and colleagues (1981) directly compared adults' and infants' performance in discriminating two speech sounds in Hindi that did not exist in English. They found that 6- to 8-month-old English-speaking infants performed similarly to Hindi-speaking adults, whereas English-speaking adults had difficulty discriminating the two sounds. This effect has been reproduced by several other studies focused on consonant perception (for a review see Werker & Curtin, 2005). Moreover, this phenomenon is not limited to consonant perception (for a review see Maurer & Werker, 2014). For example, although infants show a preference for speech sounds over non-speech sounds (Vouloumanos & Werker, 2007), this preference is broad in the sense that it includes non-speech sounds such as rhesus monkey calls as well as human speech sounds, and only narrows to human speech by the age of 3 months (Vouloumanos, Hauser, Werker, & Martin, 2010).

While perceptual narrowing was first observed with speech and auditory stimuli, it has been shown that a similar pattern is evident in the audio-visual matching of oral and visual speech. Infants at 2 months (Kuhl & Meltzoff, 1982) and 4 months (Patterson & Werker, 2003) of age have been shown to attend more to a visual stimulus that matches a speech sound stimulus, however, this ability is lost for non-native speech sounds by the end of the first year of life (Pons, Lewkowicz, Soto-Faraco, & Sebastian-Galles, 2009). Together, these observations hint towards a multimodal quality for perceptual narrowing.

1.3.1 Perceptual Narrowing and Face Perception

As previously mentioned, adults' face expertise is tuned towards better recognition of upright own-race faces. Infants have been shown to prefer looking towards face-like stimuli (for a review see Maurer & Werker, 2014), however, it has been suggested that this reflects a preference for general visual properties such as top-heaviness (Cassia, Turati, & Simion,

2004), contrast polarity (Farroni et al., 2005), and optimal range of spatial frequencies (Kleiner, 1987) rather than expertise. The basic signs of face expertise don't emerge until 3 months of age when extensive scanning of faces and fixation on internal features develop in infants (Hainline, 1978; Haith, Bergman, & Moore, 1977; Maurer & Salapatek, 1976). During the first year of life, visual attention patterns differentiate to better discriminate native and non-native stimuli hinting towards the possibility of perceptual narrowing (Maurer & Weker, 2014). Perceptual narrowing in face perception is evident first in the emergence of species specificity. Pascalis et al. (2002) showed that infants at 6 months have similar discrimination ability for both human and monkey faces but lose this ability for monkey faces at 9 months. Converging evidence from other studies using faces of Barbery macaques (Fair, Flom, Jones, & Martin, 2012; Pascalis et al., 2005) and faces of sheep (Simpson, Varga, Frick, & Fragaszy, 2011) showed that the discrimination ability for other-species faces decreases between ages 6 to 9 months, while the ability to discriminate human faces is maintained or enhanced.

Infants prefer to look at an own-race face when paired with an other-race face by the age of 3 months, however their ability to discriminate between two adult faces of own and other race is similar. Better performance for own-race face discrimination is not evident until 9 months of age. In a study conducted by Kelly et al. (2007), 3-month-old Caucasian infants were familiarized with a face followed by presentation of the same face and a novel face. Regardless of the race of the faces, 3-mounth-olds attended to the novel face more. Similarly, in a study of Chinese infants, 3-month-olds showed similar novelty preference for own- and other-race faces (Kelly et al., 2009). On the other hand, both studies showed that infants at 9 months of age showed a clear ORE—Caucasian infants attended more to Caucasian faces (Kelly et al., 2007), whereas Chinese infants attended more to Chinese

faces (Kelly et al., 2009). Additionally, in a study conducted by Ferguson et al. (2009) Caucasian infants showed holistic processing of upright own-race faces both at 4 and 8 months of age, but only at 4 months of age for other-race faces.

On the other hand, results from two older studies suggest that Caucasian infants at 3 months of age demonstrate a novelty preference for new Caucasian faces but fail to show such preference for Asian faces (Hayden, Bhatt, Joseph et al., 2007; Sangrigoli & de Schonen, 2004b). These results may weaken the evidence for perceptual narrowing. However, some important points need to be noted that help interpret these results. First, among all the 5 studies mentioned here, only the studies conducted by Kelly et al (2007 & 2009) reported results from an Asian participant group to mitigate the concerns that the Caucasian stimuli might have been more physically distinct than faces from other races used in the studies. Moreover, the older studies used black and white faces with no external features as the stimuli (Hayden, Bhatt, Joseph et al., 2007; Sangrigoli & de Schonen, 2004b), whereas the other 3 studies used coloured photos with external features (Ferguson et al., 2009; Kelly et al., 2009; Kelly et al., 2007). This is particularly important in the light of evidence showing that fixation on internal features begins to develop at the age of 3 months and studies suggesting this development have been only done on own-race faces (Hainline, 1978; Haith, Bergman, & Moore, 1977; Maurer & Salapatek, 1976). Therefore, it may be that at that age infants can rely on internal features for recognition of own-race faces but develop more advanced tools for holistic processing at a later stage.

In addition to the face discrimination studies that have been discussed so far, results from an infant face diet study lend more support to perceptual narrowing in face recognition. Sugden et al. (2014) report that infants receive more exposure to own-race faces and spend more time exclusively looking at such faces. They make the argument that perceptual

narrowing requires that infants have minimal exposure to other-race faces; and the highly exclusive exposure to own-race faces revealed by their results support this idea.

1.4 Research Question and Hypotheses

Adults' expertise in recognizing own-race faces emerges through the narrowing of face perception abilities based on differential exposure to face ethnicities during development. However, perceptual narrowing only predicts the impact of lack of exposure to a certain face type on the development of face recognition system and does not predict how the system will develop if an individual is sustainably exposed to two or more face types. Here, our purpose is to further evaluate such sustained high exposure to more than one face ethnicity and predict its impact on face recognition abilities. To examine this question, we postulate three competing hypotheses.

1.4.1 Hypothesis 1: *Resource Bottleneck*

Based on the *Resource Bottleneck* hypothesis, in order to avoid a generalist face recognition system, neural resources must be focused narrowly on a limited range of faces (e.g., on own-race faces). High exposure to a great variety of faces including multiple races will lead to a divide in allocation of resources that may hinder attainment of expertise in face recognition. Based on this hypothesis, sustained high exposure to faces of more than one ethnicity will prevent specialization and lead to a generalist, 'jack of all trades, master of none" system. Consequently, individuals exposed to faces of multiple ethnicities will show less-refined recognition for both own- and other-race faces.

This hypothesis can be traced back to a study conducted by Yurovsky, Fricker, Yu, & Smith (2014) which showed that presence of fewer visual stimuli in the visual field can facilitate learning. Indeed, one of the explanations suggested by Sugden et al., (2018) regarding the disproportionate amount of exposure to particular face types and little to no

exposure to other face types was based on this idea. The homogenous nature of infants' exposure to specific face types may stem from a purposeful avoidance of exposure to multiple types of visual stimuli, which may hinder learning.

1.4.2 Hypothesis 2: *Use/Disuse*

Based on the *Use/Disuse* hypothesis, the other-race effect is caused by a lack of exposure to, and a lack of necessity to discriminate, faces of other-races, rather than a scarcity of neural resources. This account predicts that sustained high exposure to more than one face ethnicity will lead to individuals performing at native or near-native level in recognizing both own- and other-race faces.

This account can be traced back to one of the proposed frameworks by Gottlieb (1976) put forth to explain possible impacts of perceptual experience. According to this framework, which Gottlieb refers to as "Maintenance/Loss", systems are dependent on experience for their maintenance and lack of experience can lead to loss of abilities within systems. It is worth noting that in addition to the final effect of exposure on development, Gottlieb also focuses on the timing of exposure to perceptual experience, maturation of different behavioural systems prior to exposure to experience, and in general the process through which experience shapes development. Our present hypothesis pertains only to the net impact of experience on the perceptual abilities of the mature system.

Moreover, the results from a training study conducted by Scott and Monesson (2009) suggested that the perceptual window for faces perceived as other can be held open with very minimal exposure, as long as the exposure consists of individual exemplars. Such results can be interpreted through the *Use/Disuse* model.

1.4.3 Hypothesis 3: *Facilitation*

Based on the *Facilitation* hypothesis, a richer exposure to faces leads to richer face representations. Thus, this account predicts that, sustained high exposure to more than one face-ethnicity will affect enhancements in face processing and lead to better-than-native-level performance.

This account can be traced back to Gottlieb (1976) who proposed the idea of facilitation as one possible effect of perceptual experience and defined it as "*quantitative improvement that transcends what would be achieved in the absence of experience*". Our present paradigm provides for an added level of specification such that the heterogeneous nature of the exposure results in further improvements in each component face category independently of the mere quantity of exposure.

The *Facilitation* model is consistent with the results of Balas and Saville's study (2015) where participants living in more densely populated areas performed better in face recognition.

1.4.4 Research Outline

In order to discriminate between these three alternative hypotheses, we examined the performance of individuals with sustained high exposure to both East Asian and Caucasian faces (Bi-ethnics) for face memory and perception. Since the FIE is a hallmark of face expertise, we also assessed face memory in upright and inverted orientations. Bi-ethnic participants' performance across these tasks was compared to that of Caucasian and East Asian participants, who have had sustained high exposure to own-race faces and minimal exposure to other-race faces (Mono-ethnics).

2 Methods

2.1 Participants

Twenty Caucasian participants from ages 18 to 33 (*M* = 25.05; *SD* = 4.15; 15 females), 20 Bi-ethnic participants from ages 19 to 32 (*M* = 24.45; *SD* = 3.61; 16 females), and 20 East Asian participants from ages 18 to 34 (*M* = 25.05; *SD* = 4.63; 15 females) participated in the study. The protocol was approved by the ethics review boards of the University of British Columbia and Vancouver General Hospital, and informed consent was obtained in accordance with the Declaration of Helsinki.

2.1.1 Social Exposure

All participants completed a 37-item Social Exposure Survey (Shafai & Oruc, 2018) prior to participation. The answers to this survey allowed us to assess the extent to which participants had exposure to East Asian and Caucasian faces. The inclusion criteria for Caucasian and East Asian participants were: 1) Participants should have not lived in a community with predominant exposure to other-race faces for more than 3 years; 2) Participants' self-rated exposure to own-race faces while growing up (before age 13) should have been at least a score of 4(/5) and the same score for other-race faces should have been at most 2(/5) on a scale where 1 indicates "no contact" and 5 indicates "extremely frequent and regular contact". The inclusion criteria for Bi-ethnic participants were: 1) Participants were born in an East Asian or a Caucasian family with both parents having the same race and have lived in a community with predominant exposure to both Caucasian and East Asian faces; 2) Participants' self-rated exposure to both East Asian and Caucasian faces while growing up and in the past 6 months should have been at least 4(/5).

Other items on the survey collected detailed biographical information about the participants and were used as checkpoints for consistency and accuracy of self-rated exposures.

A more detailed report of the exposure statistics is as follows: Caucasian participants' mean self-rated score to own-race faces while growing up was uniformly 5 / 5, whereas the same mean score for other-race faces was 1.7 / 5 (*SD* = 0.47). All Caucasian participants uniformly reported their close relatives to be 100% comprised of Caucasians. These reports reflect minimal exposure to East Asian faces while growing up, which was paralleled by their exposure in their adult years as well. When asked about the proportion of their close friends, Caucasian participants on average reported 84% (*SD* = 16.7) to be Caucasian and 6% (*SD* = 6.0) to be East Asian. In their larger circle of colleagues and acquaintances, on average 72.5% (*SD* = 14.1) were reported to be Caucasian, whereas 17.5% (*SD* = 16.5) were reported to be East Asian. Furthermore, their mean self-rated exposure score to own-race faces in the past 6 months was 4.9 / 5 (*SD* = 0.31), whereas same mean score for other-race faces was 2.85 / 5 (*SD* = 0.93).

East Asian participants' mean self-rated exposure score to own-race faces while growing up was uniformly 5 / 5, whereas the same mean score for other-race faces was 1.45 / 5 (*SD* = 0.51). All East Asian participants uniformly reported their close relatives to be 100% East Asian. These reports reflect minimal exposure to other-race faces as well, and were paralleled by exposure in adult years. Proportion of close friends in this group was, on average, reported to be 93.5% (*SD* = 7.45) East Asian, whereas proportion of Caucasian close friends on average was reported to be 6.5% (*SD* = 7.45). East Asian participants reported their larger circle of acquaintances and colleagues on average to be 82% (*SD* = 14.7) East Asian and 16% (*SD* = 11.0) Caucasian. Additionally, their mean self-rated

exposure to own-race faces in the past 6 months was 4.85 / 5 (*SD* = 0.37), whereas this mean score for other-race faces was 3.1 / 5 (*SD* = 1.37).

All but four Bi-ethnic participants were born in Canada or the United States (of the remaining four, 2 were born in Hong Kong and 2 were born in Taiwan). Their mean self-rated exposure score to Caucasian faces while growing up was 4.65 / 5 (*SD* = 0.49) and the same mean score for East Asian faces was 4.8 / 5 (*SD* = 0.41). They all uniformly reported their close relatives to be 100% East Asian. On average, they reported their close friends to be 64.5% (*SD* = 20.9) East Asian and 30.5% (*SD* = 19.3) Caucasian. Moreover, the average proportion of Caucasians in their larger circle of acquaintances and colleagues was 40.5% (*SD* = 17.3) and the East Asian proportion was 51.0% (*SD* = 20.0). Their mean self-rated exposure score to Caucasian faces in the past 6 months was 4.8 / 5 (*SD* = 0.41) and the same mean score for East Asian faces was 4.9 (*SD* = 0.31). In all, their exposure backgrounds reflected sustained high exposure to both Caucasian and East Asian faces.

2.1.2 Media Exposure

The Social Exposure Survey also had a section asking about hours per week spent watching TV or movies with predominantly Caucasian and East Asian casts. On average Caucasians reported 13.8 (*SD* = 10.74) hours weekly spent on TV or movies and 90.5% (*SD* = 9.31) of this time was spent on programs with predominantly own-race casts, whereas the same average time for East Asian participants was 14.4 (*SD* = 12.25) hours weekly with 53.7% (*SD* = 33.6) of this time spent on programs with predominantly own-race casts. Bi-ethnic participants on average spent 12.3 (*SD* = 11.3) hours weekly watching TV or movies and 71.3% (*SD* = 28.5) of this time was spent on programs with predominantly Caucasian casts.

2.1.3 General Intelligence

We assessed all participants' non-verbal IQ (NVIQ) using the Block Design and the Matrix Reasoning subtests of the Wechsler Abbreviated Scale of Intelligence II (WASI-II; Wechsler, 2011). NVIQ scores ranged from 93 to 134 in the Caucasian group (M = 107.65; SD = 10.0), from 81 to 128 in the Bi-ethnic group (M = 108.95; SD = 11.0), and from 100 to 121 in the East Asian group (M = 110.9; SD = 6.5). No statistically significant differences were detected between the groups' NVIQ levels (One-way ANOVA; F (2,57) = 0.6102, p = 0.55).

2.2 Experiment 1: Old/New Test

The old/new face recognition task is one of the classic measures used for evaluating the ORE in face memory (MacLin & Malpass, 2001; Meissner & Brigham, 2001; Wright, Boyd & Tredoux, 2003). In this type of test participants are presented with a series of faces in a learning phase. In the recognition phase they are presented with a mixture of novel and old faces and are asked to select the faces they had seen in the previous phase.

2.2.1 Experimental Setup

Experiment 1 was programmed in Matlab (https://www.mathworks.com) using Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and was run on a Dell XPS computer and a LG LPS LED monitor. The monitor was set to a resolution of 1280 x 720 pixels. Participants were seated 57 cm from the screen.

2.2.2 Stimuli

There were two stimulus categories in this experiment: Caucasian and East Asian faces. For each category, we had 96 different face identities, half of which were female. The

stimuli within each category covered the age range 18 to 89 years old and were grouped into 12 age groups. Each age bin covered 6 years (e.g., age bin #1: 18-23, age bin #2: 24-29, etc.) and there were 8 photos per age bin, half of which were female.

The Caucasian faces stimulus set was organized using 56 face identities from the Productive Aging Lab Face Database (Minear & Park, 2004) and 40 face identities from the FACES database (Ebner, Riediger, & Lindenberger, 2008). The width of faces on screen ranged from 6.2 cm to 8.5 cm (*M* = 7.2; *SD* = 0.48) and subtended, on average, 7.2^{o} of visual angle per face-width at viewing distance of 57 cm. The East Asian faces stimulus set was developed in our lab. Photos were taken from a distance of 60 cm by iPhone 6. The width of these faces on screen ranged from 6.6 cm to 7.7 cm (*M* = 7.0; *SD* = 0.15). All faces were front facing, with a neutral expression. Faces were converted to grayscale using Photoshop CC 2018 (http://www.adobe.com) and viewed through an oval aperture. The East Asian images, subtended, on average, 7° of visual angle per face-width at viewing distance of 57 cm.

2.2.3 Procedure

The test was done in two blocks: one with Caucasian faces and one with East Asian faces. In each participant group, half were randomly assigned to start the test with the Caucasian block and the other half started with the East Asian block. As illustrated in Figure 2.1, for each block, the participants underwent two phases: a learning phase, and a recognition phase. In the learning phase participants saw a series of 48 faces, each shown on the screen for 3 seconds with a 1-s inter-stimulus interval during which a fixation cross was presented. The 48 faces consisted of 2 female and 2 male faces randomly selected from each of the 12 age bins, and were presented to the participants in a random order.

Participants started the recognition phase immediately following the learning phase. In the recognition phase, participants completed 48 two-alternative forced-choice (2AFC) trials. In each trial, one face from the learning phase ("old") was paired with a face of the same gender randomly selected from the respective age-bin in the novel face set ("new"), the set of 48 faces not seen during the learning phase. The old/new pair of faces were presented side by side in a randomly selected order (left vs. right). The participants were asked to select the face they had seen in the learning phase by pressing the left or right keys on the computer keyboard.

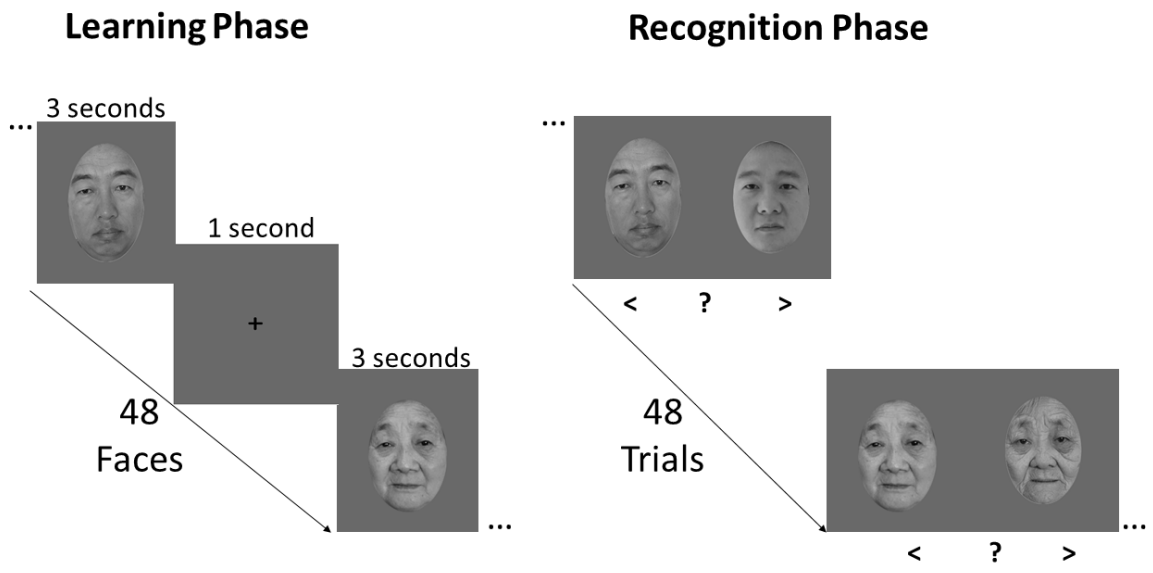

Figure 2.1 Schematic Illustration of Experiment 1 (East Asian Faces). Learning phase (on the left): participants were tasked to memorize 48 faces; each was on the screen for 3 seconds. Recognition phase (on the right): Participants completed 48 trials; each trial was a 2AFC task to indicate which one of the two faces, the one on the left or the right, was seen previously during the learning phase.

2.2.4 Data Analysis

Accuracy was calculated as the percentage of correct trials across the total number of trials (48) for each block. Accuracy scores were submitted to a repeated-measures ANOVA with stimulus category (Caucasian, East Asian) as a within-subjects factor and participant group (Caucasian, Bi-Ethnic, and East Asian) as a between-subjects factor. Posthoc pairwise comparisons were based on Tukey-Kramer tests. Accuracies were combined to yield scores for own-race (Caucasian participants for Caucasian faces and East Asian participants for East Asian faces), other-race (Caucasian participants for East Asian faces and East Asian participants for Caucasian faces), and Bi-ethic exposure. Combined scores were submitted to a one-way ANOVA with exposure type (own-race, other-race, bi-ethnic) as the between-subject factor followed by a two-tailed 2 sample t-test to evaluate the *apriori* hypothesis for bi-ethnic differences.

2.3 Experiment 2: Cambridge Face and Car Memory Tests

The Cambridge Face Memory Test (CFMT; Duchaine and Nakayama, 2006) is a measure of face memory and among the most widely used test in the face recognition literature at present (Robotham & Starrfelt, 2018). As illustrated in Figure 2.2, the test involves three phases. In the learning phase, participants learn six target face identities. Each target is presented from three different views. Immediately after this presentation, the participants complete three trials in which the individual target is presented with two other distractors. At the end of this phase the participants are presented with all the six identities from a frontal view, followed by the no-noise test phase. In this phase participants complete 30 three-alternative forced-choice (3AFC) trials where they are presented with one of the targets alongside two distractors. In the last phase the participants complete 24 trials with the same format but with visual noise added to the photos. The test has been adapted in

many ways. In order to evaluate the other-race effect McKone et al. (2012) developed *CFMT Chinese* using the same paradigm as the original test but with Chinese faces. Moreover, Dennett et al. (2012) developed the *Cambridge Car Memory Test* (CCMT), an object recognition test matched in its format with CFMT. Additionally, the inverted formats of these tests have been used to measure object and face inversion effects. In Experiment 2, we used the CFMT, the CFMT Chinese, and the CCMT, both in the upright and inverted formats.

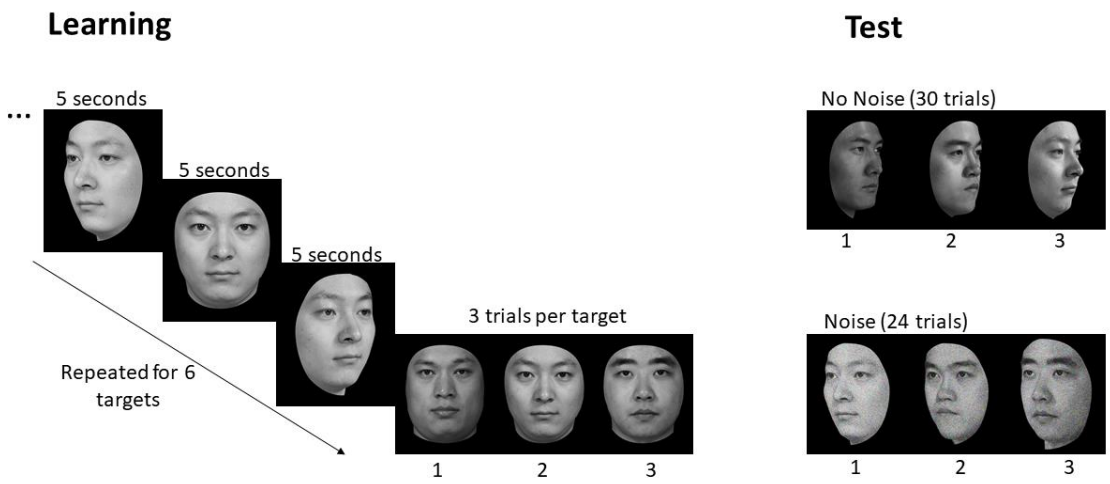

Figure 2.2 Schematic Illustration of Experiment 2 (Upright Chinese Faces). Each Cambridge memory test has 6 target stimuli. In the learning phase (on the left) participants are familiarized with those stimuli and in the test phase they are tested for their memory for the targets (on the right).

2.3.1 Experimental Setup and Stimuli

All three tests were run on a Dell XPS computer and a LG LPS LED monitor. The CFMT and the CCMT were administered using Java (TM) Platform SE binary. The CFMT Chinese was run using Matlab (https://www.mathworks.com). When running the face tests the monitor was set to a resolution of 1280 x 720 pixels. For the CCMT the monitor was set to a resolution of 1600 x 900 pixels. Participants were seated 57 cm from the screen for all

three tests. At this viewing distance, faces subtended 5.5^{o} of visual angle per face-width and Cars subtended 13° of visual angle per car-width.

2.3.2 Procedure

Half of the participants in each group were randomly assigned to start with the car test and the other half were assigned to start with faces. Within the face block, half of the participants were randomly assigned to start with Caucasian faces and the other half started with East Asian faces. For all tests, participants completed the upright task first, followed by the inverted format.

2.3.3 Data Analysis

Accuracy for each test was measured as the percentage of correct trials across the total of 72 trials. Accuracy scores were submitted to a repeated-measures ANOVA with stimulus category (Caucasian faces, East Asian faces, and Cars) as a within-subjects factor and participant group (Caucasian, Bi-Ethnic, and East Asian) as a between-subjects factor. Posthoc pairwise comparisons were based on Tukey-Kramer tests. A face inversion effect (FIE) score was computed by subtracting the inverted condition accuracy scores from those of the respective upright scores. An Alternative FIE score was also computed based on residuals in a linear regression analysis in which the inverted scores were partialled out of the upright scores in each condition. Upright accuracies and FIE scores were combined to yield scores for own-race (Caucasian participants for Caucasian faces and East Asian participants for East Asian faces), other-race (Caucasian participants for East Asian faces and East Asian participants for Caucasian faces), and Bi-ethic exposure. Combined scores were submitted to a one-way ANOVA with exposure type (own-race, other-race, bi-ethnic)

as the between-subject factor followed by a two-tailed 2 sample t-test to evaluate the *apriori* hypothesis for bi-ethnic differences.

2.4 Experiment 3: Identity Discrimination Task

As discussed earlier in the introduction, the ORE is not limited to face memory and can be observed in face discrimination tasks as well. In our study, Experiments 1 and 2 focused on face memory and recognition abilities. Here, we designed a third experiment with minimal memory demands to evaluate perceptual discrimination abilities. We chose 4 pairs of female and male, Caucasian and East Asian familiar celebrity identities and generated different morph levels for each pair. The aim of the test was to evaluate identity discrimination abilities for own- and other-race faces by finding discrimination thresholds for each pair using an adaptive psychophysical staircase centered around the morph level that represents the category boundary for the two identities. In order to find the category boundary for each participant, we evaluated the point of subjective equality (PSE) for each participant separately and used it in their respective discrimination tests.

2.4.1 Experimental Setup

This experiment was run on a computer equipped with a Cambridge Research Systems (CRS) VSG 2/3 graphics card and a SONY Trinitron 17 in. monitor (model Multiscan17seII). It was programmed in Matlab (https://www.mathworks.com) using the CRS Toolbox for Matlab and Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Participants were seated 70 cm away from the screen.

2.4.2 Stimuli

High resolution face images of 8 celebrity identities were obtained from various online resources. These were scaled to 1024 x 1024 pixels and converted to grayscale

using Photoshop CC 2018 (http://www.adobe.com). The images subtended, on average, 6.5° of visual angle per face-width at viewing distance of 70 cm. This face size has been selected to be in the range that optimizes face perception (Yang, Shafai & Oruc, 2014).

These images represented four pairs of identities: East Asian female, East Asian male, Caucasian female, and Caucasian male. For each pair, we created a morph series between the two identities using FantaMorph 4 (http://www.fantamorph.com). Each series consisted of 101 morphed images, that gradually blended the two identities in 1% steps. Some sample images from each morph series is presented in Figure 2.3.

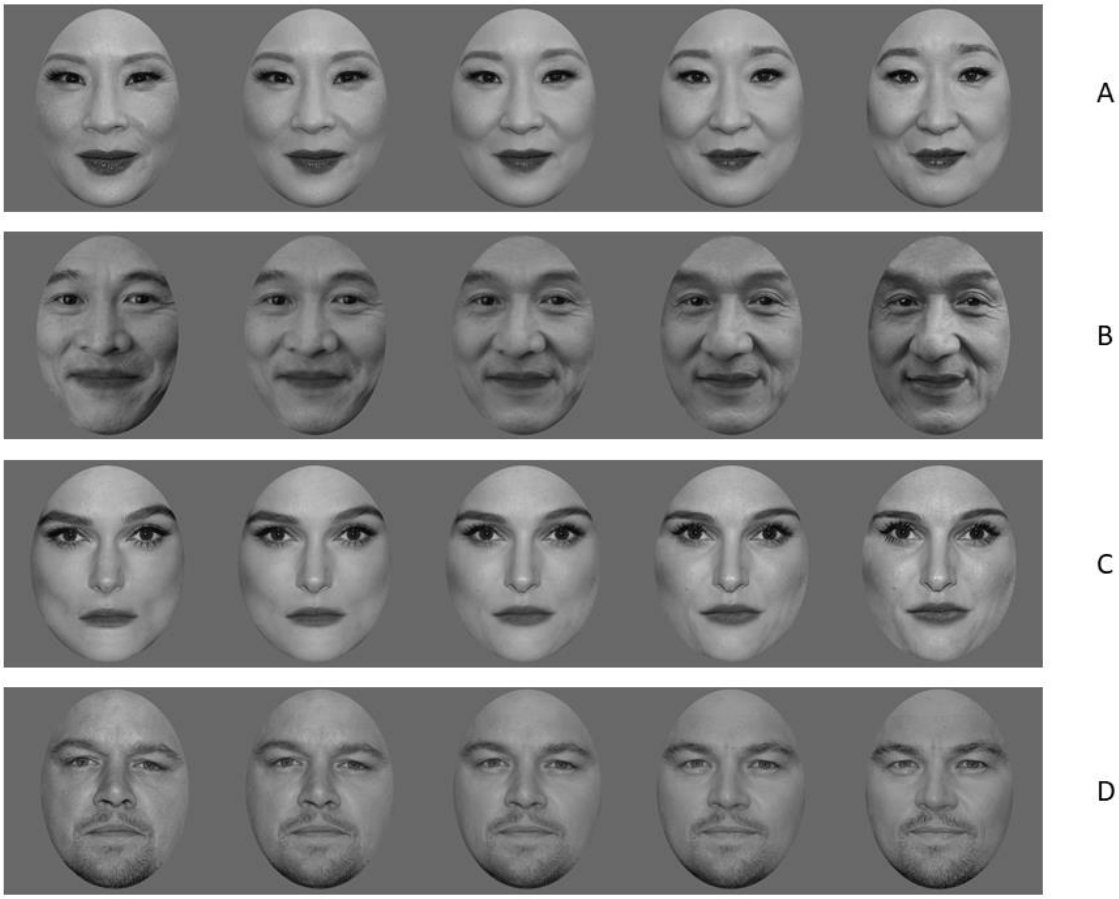

Figure 2.3 Sample Celebrity Morph Series Examples of 0%, 25%, 50%, 75%, and 100% celebrity identity morphs. (A) East Asian female: Lucy Liu (left) and Sandra Oh (right), (B) East Asian male: Jet Li (left) and Jackie Chan (right), (C) Caucasian female: Keira Knightley (left) and Natalie Portman (right), (D) Caucasian male: Matt Damon (left) and Leonardo DiCaprio (right).

2.4.3 Procedure

This experiment included four blocks: East Asian female, East Asian male, Caucasian female, Caucasian male. In each participant group, half of the participants were randomly assigned to start with Caucasian faces. For each face ethnicity, half of the participants started with the male pair. As demonstrated in Figure 2.4, each block consisted of two separate phases. The aim of the first phase was to find the point of subjective equality (PSE) for each participant. In this experiment, the PSE is defined as the morph level that lies at the perceptual category boundary of the two identities. Morph levels below the PSE would be perceived more often as identity 1, and morph levels above the PSE would more often be perceived as identity 2, and at the PSE, the participant would identify the morph as either of the two identities in the particular pair in half of the instances. In each trial, participants were presented with a fixation screen for 150 ms followed by a screen showing one of the morph levels for 500 ms, followed by a choice screen displaying the two identities. Participants pressed the “1” key or the “2” key to select the identity they thought the image they had seen had the most resemblance to. A psychophysical staircase was implemented to control morph level decrements or increments in each trial using two randomly interleaved staircases of 40 trials. QUEST procedure (Watson & Pelli, 1983) in Psychophysics toolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) was used to execute the staircases. Criterion accuracy was set to 50% so that the threshold estimate represents the PSE. Blocks where one of the PSE estimates was more than 20% larger than the other were repeated until reliable thresholds were obtained.

The aim of the second phase was to find identity discrimination thresholds, i.e., the smallest morph difference around the PSE that allows successful discrimination of the two identities. This procedure was modeled after the expression discrimination paradigm of

Oruc, Shafai, and Iarocci (2018), which in turn was modeled after a classical orientation discrimination paradigm (e.g., Regan & Beverley, 1985). In each trial, participants were presented with a fixation screen for 150 ms followed by two 500-ms stimulus displays, each showing a distinct morph level, equidistance from the PSE towards each identity. The participants were instructed to select which one of the two stimuli presented to them resembled one of the identities in that pair by pressing the “1” key or the “2” key. Half of the participants completed the task with one of the identities in each pair as the target identity with the other half having the second identity as the target. In each trial, the morph difference between the two stimuli was controlled by a psychophysical staircase with an 82% criterion accuracy. Two independent estimates of the discrimination threshold were obtained via two randomly interleaved staircases of 40 trials each. The staircases were based on the QUEST procedure (Watson & Pelli, 1983) implemented in the Psychophysics toolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). Auditory feedback was provided in the form of a single click for correct responses, and two clicks for incorrect responses. Blocks where one of the threshold estimates was more than 20% higher than the other were repeated until reliable thresholds were obtained.

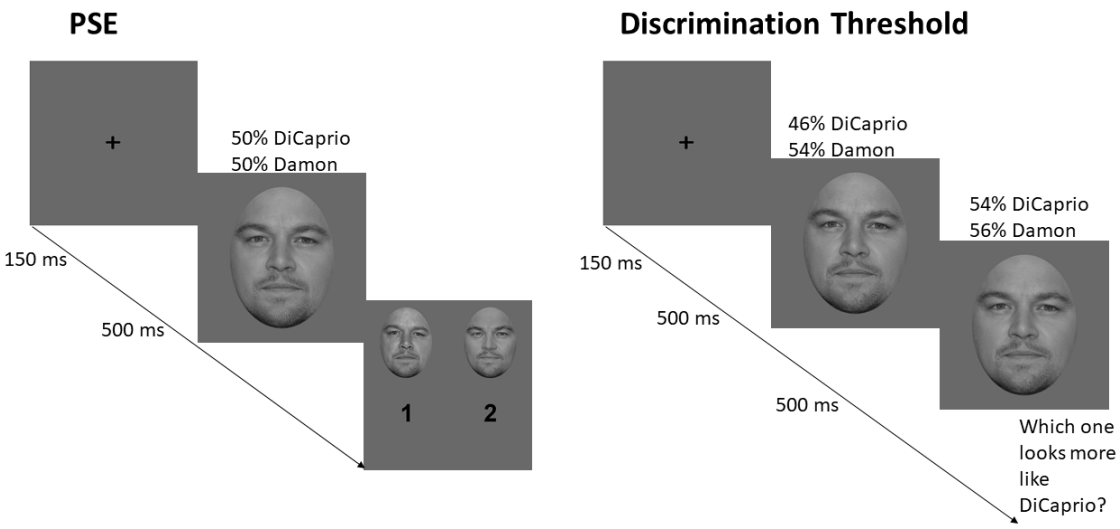

Figure 2.4 Schematic Illustration of Experiment 3 (Caucasian Male Pair). This experiment involved two phases: 1) PSE phase (left): in each trial participants were presented with one of the morph levels and asked to chose which of the two identities the presented photo resembled more.; 2) Discrimination phase (right): in each trial participants were presented with 2 different morph levels around the PSE and were asked to chose which of the 2 resembled one of the identities more.

2.4.4 Data Analysis

For each stimulus block, discrimination thresholds were measured at 82% criterion accuracy in a 2AFC paradigm. The two independent estimates of the PSE were averaged to obtain a final PSE value for each participant and each stimulus category. The participant-specific PSE values were incorporated individually into the discrimination phase of the experiment, such that identity discrimination thresholds were measured around each participant's individual PSE value. The two independent measurements of the identity discrimination thresholds were averaged to obtain a final threshold estimate. Discrimination thresholds were submitted to a repeated-measured ANOVA with stimulus category (Caucasian and East Asian) as a within-subjects factor and participant group (Caucasian, Bi-Ethnic, and East Asian) as a between-subjects factor. Posthoc pairwise comparisons

were based on Tukey-Kramer tests. Thresholds from the Caucasian and East Asian tasks were combined to yield scores for own-race (Caucasian participants for Caucasian faces and East Asian participants for East Asian faces), other-race (Caucasian participants for East Asian faces and East Asian participants for Caucasian faces), and Bi-ethic exposure. Combined thresholds were submitted to a one-way ANOVA with exposure type (own-race, other-race, bi-ethnic) as the between-subject factor followed by a two-tailed 2 sample t-test to evaluate the *apriori* hypothesis for bi-ethnic differences.

3 Results

3.1 Experiment 1: Old/New Test

3.1.1 General Analysis

Figure 3.1 shows accuracy for East Asian and Caucasian face recognition in the old/new test among the three participant groups. Accuracies for each stimulus category for each participant were calculated as the percentage of the number of correct trials with respect to the total number of 2AFC trials completed. Accuracies were submitted to a repeated-measures ANOVA with stimulus category (East Asian and Caucasian) as the within-subjects factor and participant ethnicity (Caucasian, Bi-ethnic, and East Asian) as the between-subjects factor. There were no significant main effects for stimulus category (F (1,57) = 0.00, p = .949) or participant ethnicity (F (2,57) = 0.08, p = .920). However, there was a significant interaction between stimulus category and participant ethnicity (F (2,57) = 13.55, p < .001, η_p^2 = .322). Caucasian and East Asian participants performed better for own-race faces compared to other-race. Posthoc multiple comparisons showed that Caucasian participants' accuracy (M = 76.46%) was significantly higher than that of East Asian participants (M = 69.90%) for the Caucasian stimulus set (Tukey-Kramer, p = .011, d = 0.69). Conversely, East Asian participants' accuracy (M = 76.25%) was significantly higher than Caucasian participants (M = 69.17%) for the East Asian stimuli (Tukey-Kramer, p = .004, d = 0.68). Bi-ethnic participants performed at near-native levels for East Asian (M = 74.48%) and Caucasian (M =73.33%) faces, and posthoc comparisons of their accuracies were not significantly different from Caucasian and East Asian participants for the East Asian stimuli (Tukey-Kramer, p = .07 and p = .93 respectively) as well as the Caucasian stimuli (Tukey-Kramer, p = .56 and p = .45 respectively).

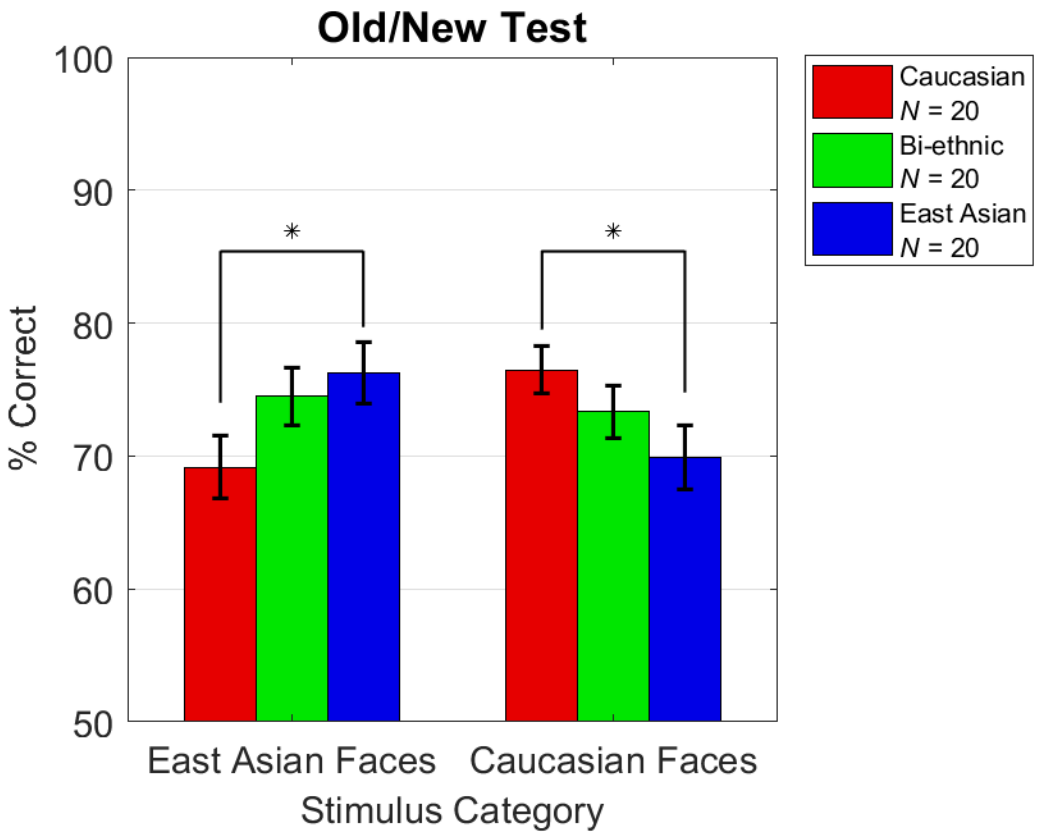

Figure 3.1 Experiment 1 Results: General Analysis. Average accuracies for East Asian faces (left) and Caucasian faces (right) across the three participant groups. Error bars represent standard errors (+/- 1SE).

3.1.2 Own-Race vs Other-Race vs Bi-ethnic Analysis

Figure 3.2 shows accuracy for own-race (Caucasian participants for Caucasian faces and East Asian participants for East Asian faces), other-race (Caucasian participants for East Asian faces and East Asian participants for Caucasian faces), and Bi-ethic exposure. Accuracies were submitted to a one-way ANOVA with exposure type (own-race, other-race, bi-ethnic) as the between-subject factor. There was a significant main effect of exposure ($F(2,117)= 5.123$, $p = .007$, $\eta_p^2 = .081$). Further analysis using two-tailed 2 sample t-test showed a significant difference between own-race and other-race exposure ($t(78) = 3.095$, $p = .0027$, $d = 0.69$), a near-significant difference between bi-ethnic and other-race

exposure ($t(78) = 1.964$, $p = .053$, $d = 0.44$), and no significant difference between own-race and bi-ethnic exposure ($t(78) = 1.198$, $p = .234$).

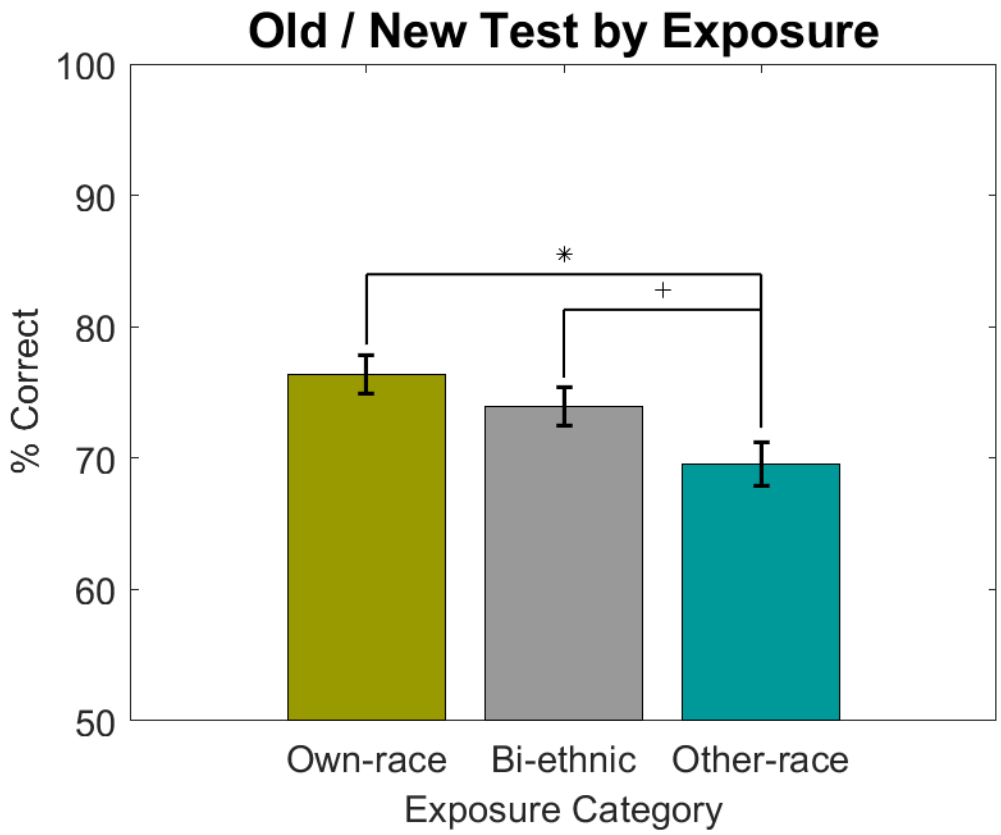

Figure 3.2 Experiment 1 Results: Own-race vs Other-race vs Bi-ethnic Analysis. Average accuracies for own-race (left), bi-ethnic (middle), and other-race (right) exposure for the old/new task. Error bars represent standard errors (+/- 1SE).

3.2 Experiment 2: Cambridge Memory Tests

3.2.1 Upright Accuracies: General Analysis

Figure 3.3 shows results from the upright Cambridge memory tests. Accuracies for the upright tests were measured as the percentage of correct trials and were submitted to a repeated-measures ANOVA with stimulus category (Caucasian faces, East Asian faces, and Cars) as the within-subjects factor and participant ethnicity as the between-subjects

factor. Results revealed a significant main effect for stimulus category (F (2,57) = 64.35, p << .001, η_p^2 = .53). In general, faces were recognized better than cars and all participants performed similarly for the cars task. Posthoc multiple comparisons showed that both East Asian faces and Caucasian faces were recognized significantly better than Cars (Tukey-Kramer, both ps << 0.001). No significant main effect for participant ethnicity was present (F (2,57) = 1.82, p = .17), however a significant interaction between participant ethnicity and stimulus category was observed (F (4,114) = 8.34, p << .001, η_p^2 = .226). Similar to the results of Experiment 1, Caucasian and East Asian participants performed better for own-race faces compared to each other. Posthoc analysis of the interaction between participants' and stimulus ethnicity showed that East Asian participants' accuracy for recognizing East Asian faces (M = 89.79%) was significantly higher than that of Caucasian participants (M = 80.76%) for East Asian faces (Tukey-Kramer, p = .028, d = 0.98). Conversely, Caucasian participants' accuracy for recognizing Caucasian faces (M = 88.19%) was significantly higher than that of East Asian participants (M = 76.39%) (Tukey-Kramer, p = .0005, d = 1.50). Bi-ethnic participants showed near-native level performance in recognizing East Asian and Caucasian faces. Bi-ethnic participants' accuracies for Caucasian faces (M = 87.77%) was significantly greater than that of East Asian participants (Tukey-Kramer, p = .0018, d = 1.39) and their accuracy for recognizing East Asian faces (M = 89.23%) was higher than Caucasian participants at a near-significant level (Tukey-Kramer, p = .052, d = 0.86).

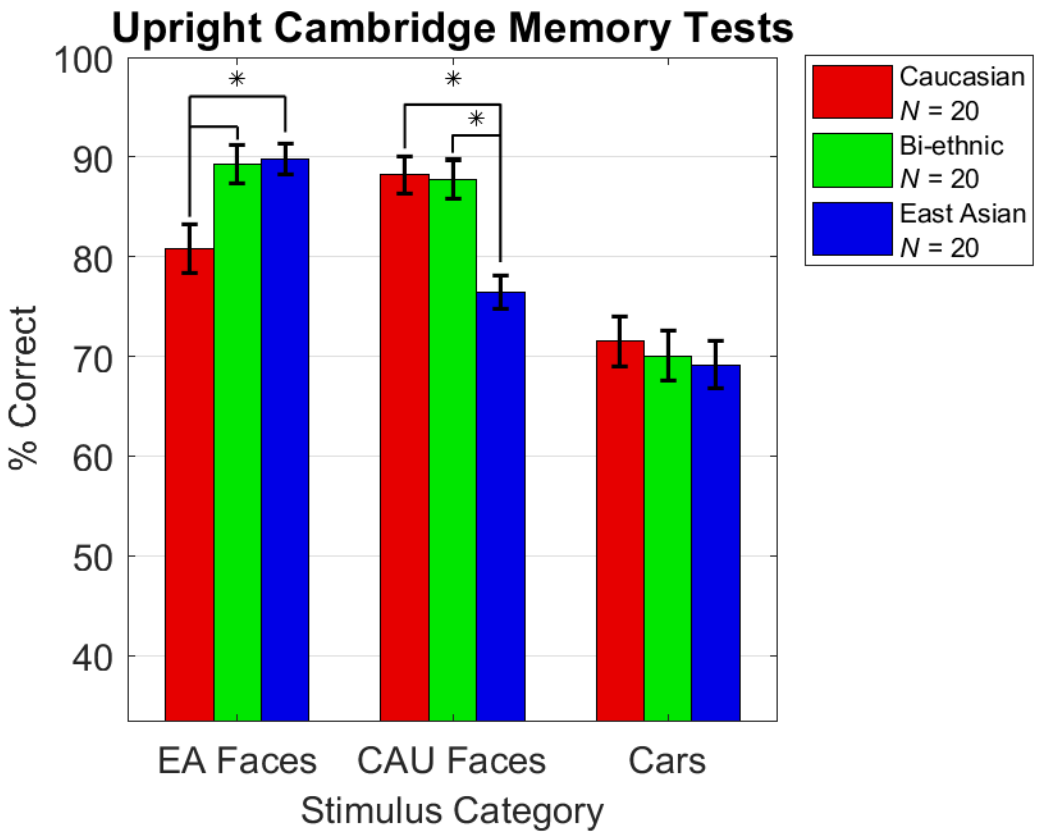

Figure 3.3 Results from Upright Cambridge Memory Tests: General Analysis. Average accuracies from CFMT Chinese (left), CFMT (middle), and CCMT (right) across the three participant groups. Error bars represent +/- 1SE.

3.2.2 Upright Accuracies: Own-race vs Other-race vs Bi-ethnic Analysis

Figure 3.4 shows CFMT scores for own-race (Caucasian participants for Caucasian faces and East Asian participants for East Asian faces), other-race (Caucasian participants for East Asian faces and East Asian participants for Caucasian faces), and Bi-ethic exposure. Scores were submitted to a one-way ANOVA with exposure type (own-race, other-race, bi-ethnic) as the between-subject factor. There was a significant main effect of exposure type ($F(2,117)= 18.573$, $p << .001$, $\eta_p^2 = .241$). Further analysis using two-tailed 2 sample t-test showed significant differences between own-race and other-race exposure ($t(78) = 5.414$, $p << .0001$, $d = 1.21$) as well as bi-ethnic and other-race exposure ($t(78) =$

4.782, $p << .0001$, $d = 1.09$) and no significant difference between own-race and bi-ethnic exposure ($t(78) = 0.268$, $p = .790$).

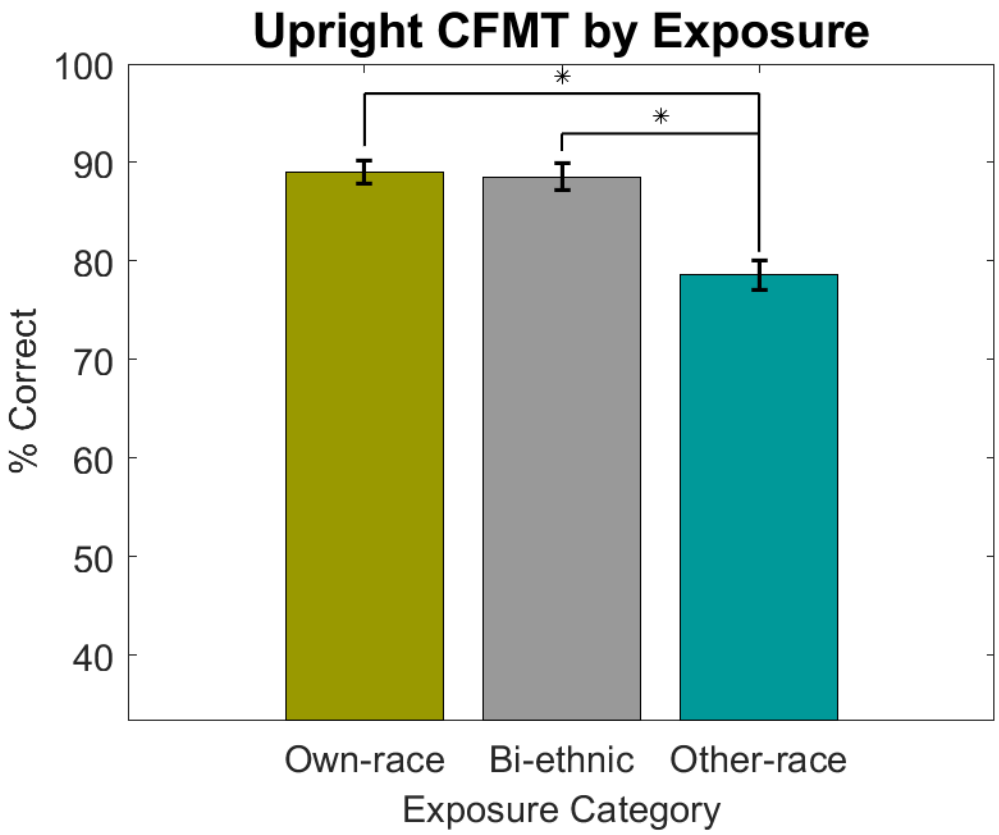

Figure 3.4 Results from Upright CFMT: Own-race vs Other-race vs Bi-ethnic Analysis. Average accuracies for own-race (left), bi-ethnic (middle), and other-race (right) exposure for the CFMT. Error bars represent standard errors (+/- 1SE).

3.2.3 Inverted Accuracies

Results of the inverted Cambridge tests are plotted in Figure 3.5. Accuracies were submitted to a repeated-measures ANOVA with stimulus category (Caucasian faces, East Asian faces, and Cars) as the within-subjects factor and participant ethnicity (Caucasian, East Asian, and Bi-ethnic) as the between-subjects factor. No significant main effect for participant ethnicity (F (2,57) = 0.26, $p = .778$) was observed. A significant main effect of

stimulus category was detected (F (2,118) = 25.43, $p << .001$, $\eta_p^2 = .309$). Participants performed better at recognizing inverted cars compared to faces. No ORE for faces was present as indicated by a lack of significant interaction between participant ethnicity and stimulus category (F (4,114) = 1.56, $p = .19$). Posthoc comparisons showed a significant difference between inverted car accuracies (M = 69.56 %) and inverted East Asian face (M = 58.61%) and inverted Caucasian face (M = 63.75%) accuracies (Tukey-Kramer, $p << .001$, $p = .0005$ respectively).

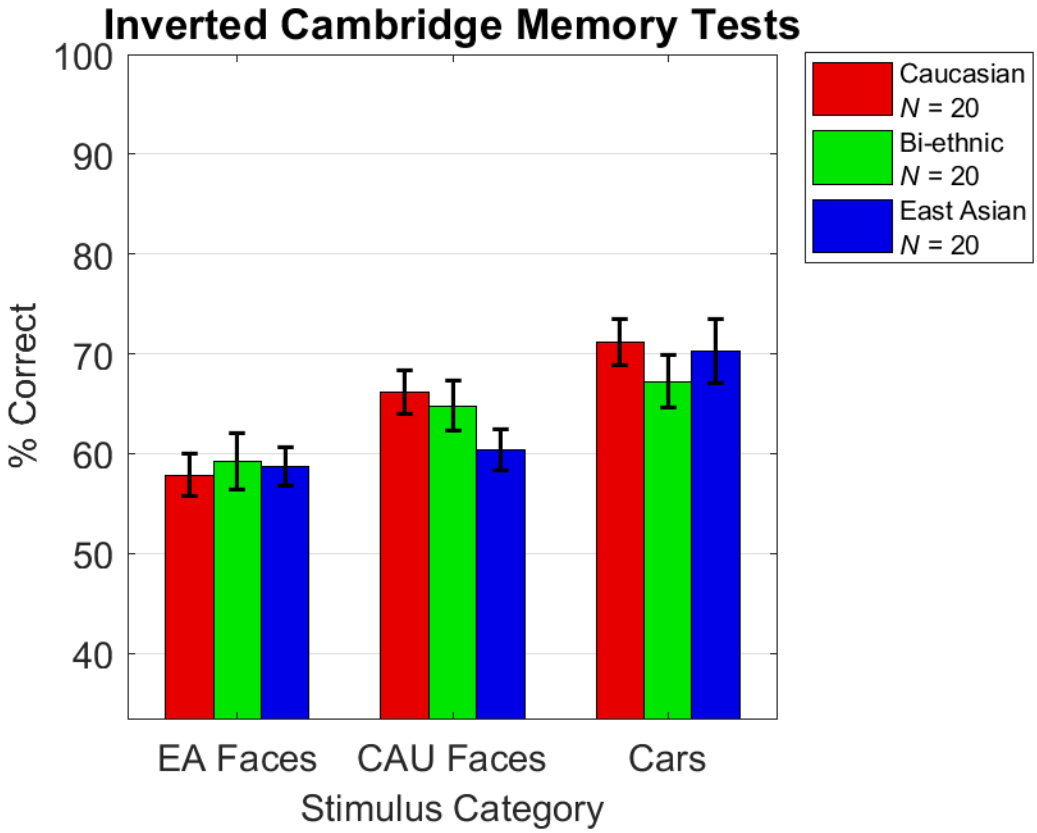

Figure 3.5 Results of the Inverted Cambridge Memory Tests. Average accuracies for the inverted CFMT Chinese (left), inverted CFMT (middle), and inverted CCMT (right) are plotted across the three participant groups. Error bars represent +/- 1SE.

3.2.4 FIE – Subtraction Scores: General Analysis

Figure 3.6 shows the difference scores between upright and inverted versions for all Cambridge tests across the three participant groups. Here, the difference scores were calculated as a measure of the FIE and were submitted to a repeated-measures ANOVA with stimulus category (Caucasian faces, East Asian faces, and cars) as the within-subjects factor and participant ethnicity (Caucasian, East Asian, and Bi-ethnic) as the between-subjects factor. Results of this test revealed a significant main effect of stimulus category (F (2,57) = 126.86, $p << .001$, $\eta_p^2 = .69$) and a significant interaction between participant ethnicity and stimulus category (F (4,114) = 2.802, $p = .03$, $\eta_p^2 = .09$), but no significant main effect for participant ethnicity (F (2,57) = 1.30, $p = .281$). The inversion effects for faces were larger than cars. Posthoc pair-wise comparisons showed that the inversion effect for cars (M = 0.69) was significantly lower than the FIE for Caucasian (M = 20.37) and East Asian (M = 27.99) faces with the significance levels from Tukey-Kramer test being $p << .001$ in both comparisons (d = 1.76 and 2.43 respectively). Own-race FIE scores were larger for Caucasian and East Asian participants compared to each other. FIE scores for Caucasian faces were numerically higher in the Caucasian group (M = 22.08) compared to the East Asian group (M = 16.04), but this difference was not statistically significant (Tukey-Kramer, $p = .55$, d = 0.66). As well, FIE scores for East Asian faces were numerically, but not significantly, larger in the East Asian group (M = 31.04) than the Caucasian group (M = 22.92) (Tukey-Kramer, $p = .16$, d = 0.84). Bi-ethnic participants' FIE scores for Caucasian (M = 22.98) and East Asian (M = 30.0) faces were similar to own-race FIE scores for Mono-ethnic groups. Their FIE scores were not significantly different from Caucasians for Caucasian faces (Tukey-Kramer, p = 1) and from East Asians for East Asian faces (Tukey-Kramer, p = 1).

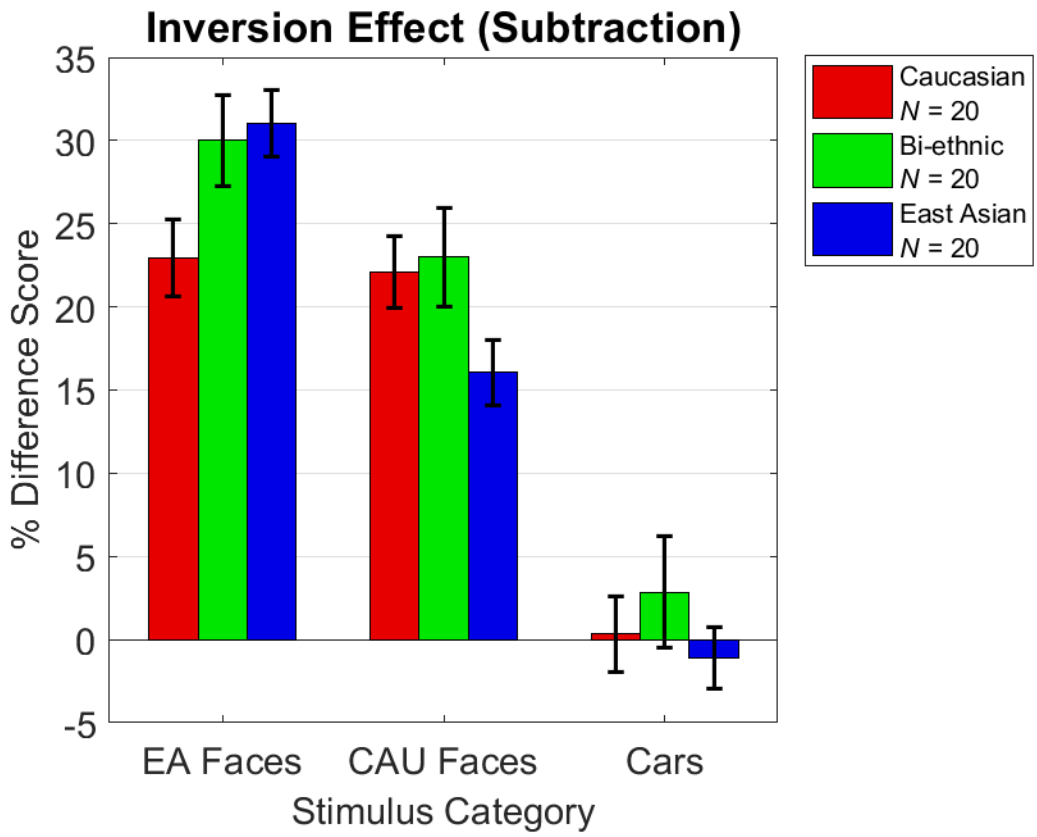

Figure 3.6 Face Inversion Effect (Subtraction Scores): General Analysis. The difference between upright and inverted scores in the Cambridge memory tests were used to index face inversion effect (FIE) across all stimulus categories. Error bars represent +/- 1SE.

3.2.5 FIE – Subtraction Scores: Own-race vs Other-race vs Bi-ethnic Analysis

Figure 3.7 shows FIE scores from CFMTs for own-race (Caucasian participants for Caucasian faces and East Asian participants for East Asian faces), other-race (Caucasian participants for East Asian faces and East Asian participants for Caucasian faces), and Bi-ethic exposure. Scores were submitted to a one-way ANOVA with exposure type (own-race, other-race, bi-ethnic) as the between-subject factor. There was a significant main effect of exposure type ($F(2,117)= 5.26$, $p = .0065$, $\eta_p^2 = .082$). Further analysis using two-tailed 2

sample t-test showed significant differences between own-race and other-race exposure ($t(78) = 3.123$, $p = .0025$, $d = 0.7$) as well as bi-ethnic and other-race exposure ($t(78) = 2.687$, $p = .0088$, $d = 0.6$) and no significant difference between own-race and bi-ethnic exposure ($t(78) = 0.026$, $p = .979$).

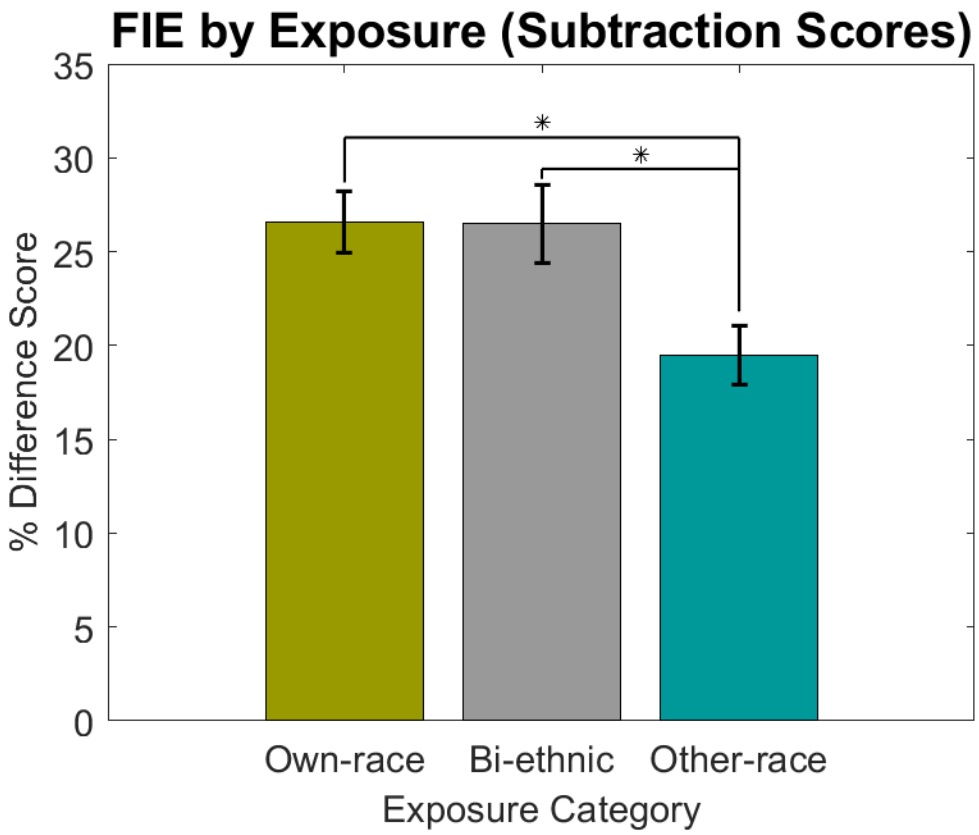

Figure 3.7 Face Inversion Effect (Subtraction Scores): Own-race vs Other-race vs Bi-ethnic Analysis. Average FIE scores for own-race (left), bi-ethnic (middle), and other-race (right) exposure. Error bars represent standard errors (+/- 1SE).

3.2.6 FIE – Residuals: General Analysis

The difference between upright vs. inverted accuracies has commonly been used to derive a FIE score (Hancock & Rhodes, 2008; Rhodes, Hayward, & Winkler, 2006). A limitation of this approach is that the subtraction scores obscure the relative contributions

from the two components. In other words, by merely evaluating the difference, it is unclear whether a higher inversion effect in one condition vs. another is rooted in a higher accuracy in the upright orientation, a lower accuracy in the inverted orientation, or a combination of both. An alternative method has been proposed by DeGutis et al. (2013) in the context of assessing holistic processing via the part-whole effect (Tanaka & Farah, 1993), and the composite face effect (Young, Hellawell, & Hay, 2013). In this linear regression-based alternative, the main outcome measure is based on the residuals that remain after the control condition is regressed out from the condition of interest. We utilize this approach in our present study by regressing out the inverted accuracies from those of the upright. This allows us to remove the inverted condition variance, yielding a FIE score that is uncorrelated with performance in the inverted orientation.

In Figure 3.8 the top row shows scatter plots of all participants' accuracies for the upright condition versus the inverted condition for each stimulus category around an identity line. The bottom row shows the same data with the best linear fit superimposed and how the residual scores would differ in this case compared to the subtraction methods.

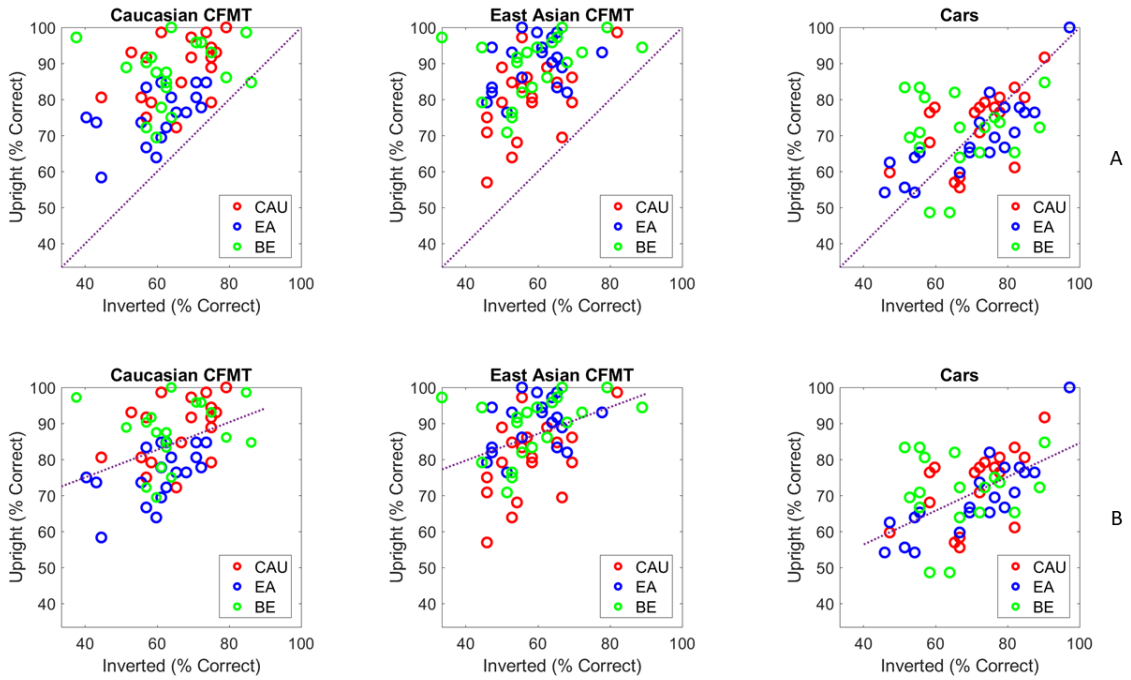

Figure 3.8 Upright vs Inverted Scores across three Cambridge Tests. A scatterplot of all participants' upright vs inverted scores across the three tests superimposed with the identity line (A) and the best fitting line (B).

The FIE scores were obtained with the regression-based approach described in DeGutis et al. (2013) based on the residuals. Average FEI scores for each participant group and stimulus category are shown in Figure 3.9. In contrast to the subtraction-based FIE scores presented in Section 3.2.4, here, a score of zero does not indicate a lack of difference between the upright and inverted scores. Instead, a score of zero indicates that the upright accuracies are commensurate with the inverted accuracies based on the performance of the entire participant group. FIE scores are positive when upright accuracies exceed what would be commensurate with the corresponding inverted performance, and negative if they fall short of them.

Residuals were submitted to a repeated-measures ANOVA with stimulus category (Caucasian faces, East Asian faces, and Cars) as the within-subjects factor and participant

ethnicity (Caucasian, East Asian, and Bi-ethnic) as the between-subjects factor. The analysis did not show a significant main effect for participant ethnicity (F (2,57) = 2.51, p = .09) or stimulus category (F (2,54) = 0.00, p = 1), but revealed a significant interaction between stimulus category and participant ethnicity (F (4, 114) = 7.04, $p << 0.001$, η_p^2 = .198). Caucasian and East Asian participants showed higher FIE scores for own-race faces compared to each other. Posthoc pair-wise comparisons showed that Caucasian participants' FIE scores for Caucasian faces (M = 3.17) were significantly higher than the same for East Asian participants (M = -6.43) (Tukey-Kramer, p = .006, d = 1.37). Conversely, East Asian participants' FIE scores for East Asian faces (M = 3.14) were significantly higher than that for Caucasian participants (M = -5.55) (Tukey-Kramer, p = .02, d = 1.05). Interestingly, Bi-ethnic participants residual FIE scores showed an ORE comparable to Mono-ethnic participants' scores for own-race faces and higher OREs than that of Mono-ethnics for other-race faces. Bi-ethnic participants' FIE scores for Caucasian faces (M = 3.26) were significantly higher than the same for East Asian participants (Tukey-Kramer, p = .005, d = 1.21) and their FIE scores of for East Asian faces (M = 2.41) were significantly higher than that of Caucasian participants (Tukey-Kramer, p = .048, d = 0.88). Bi-ethnic participants' FIE scores did not differ from own-race FIE scores for either of the Mono-ethnic groups (p = 1 for both comparisons).

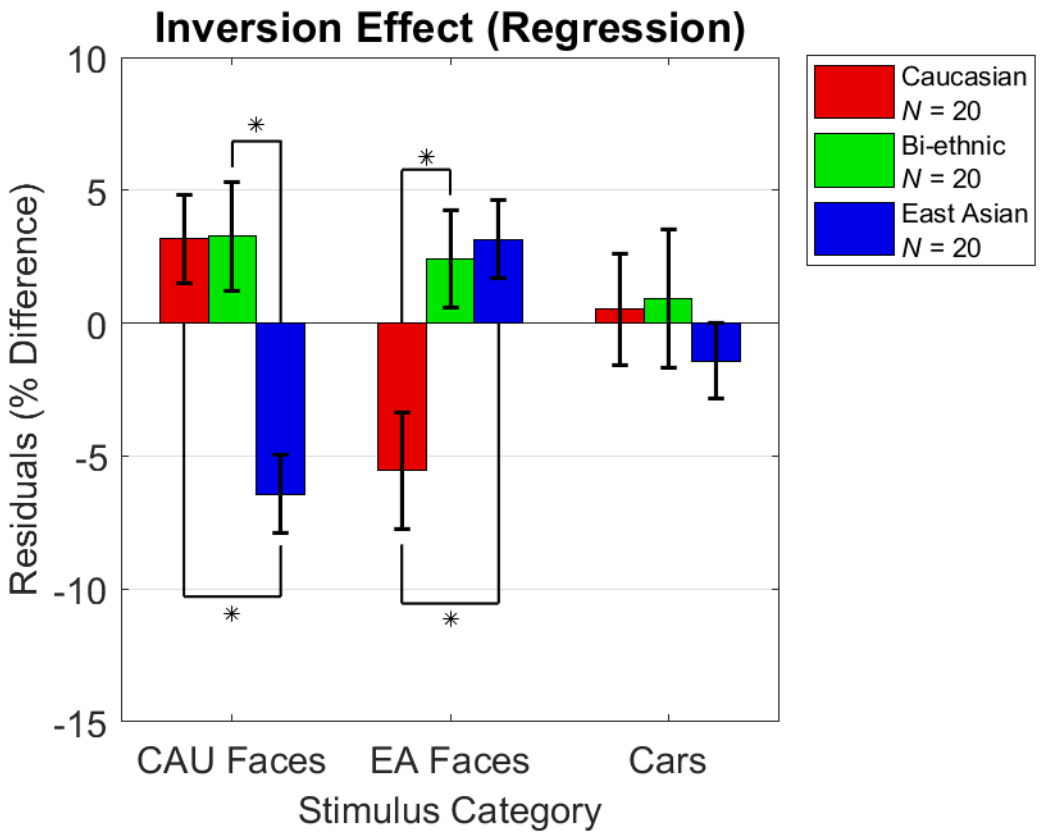

Figure 3.9 Face Inversion Effect (Regression): General Analysis. Residual scores as a second index for FIE across all stimulus categories. Error bars represent +/- 1SE.

3.2.7 FIE – Residuals: Own-race vs Other-race vs Bi-ethnic Analysis

Figure 3.10 shows residual scores from CFMTs for own-race (Caucasian participants for Caucasian faces and East Asian participants for East Asian faces), other-race (Caucasian participants for East Asian faces and East Asian participants for Caucasian faces), and Bi-ethic exposure. Scores were submitted to a one-way ANOVA with exposure type (own-race, other-race, bi-ethnic) as the between-subject factor. There was a significant main effect of exposure type ($F(2,117)= 17.063$, $p << .001$, $\eta_p^2 = .226$). Further analysis using two-tailed 2 sample t-test showed significant differences between own-race and other-race exposure ($t(78) = 5.389$, $p << .001$, $d = 1.2$) as well as bi-ethnic and other-race

exposure ($t(78) = 4.691$, $p << .001$, $d = 1.05$) and no significant difference between own-race and bi-ethnic exposure ($t(78) = 0.185$, $p = .853$).

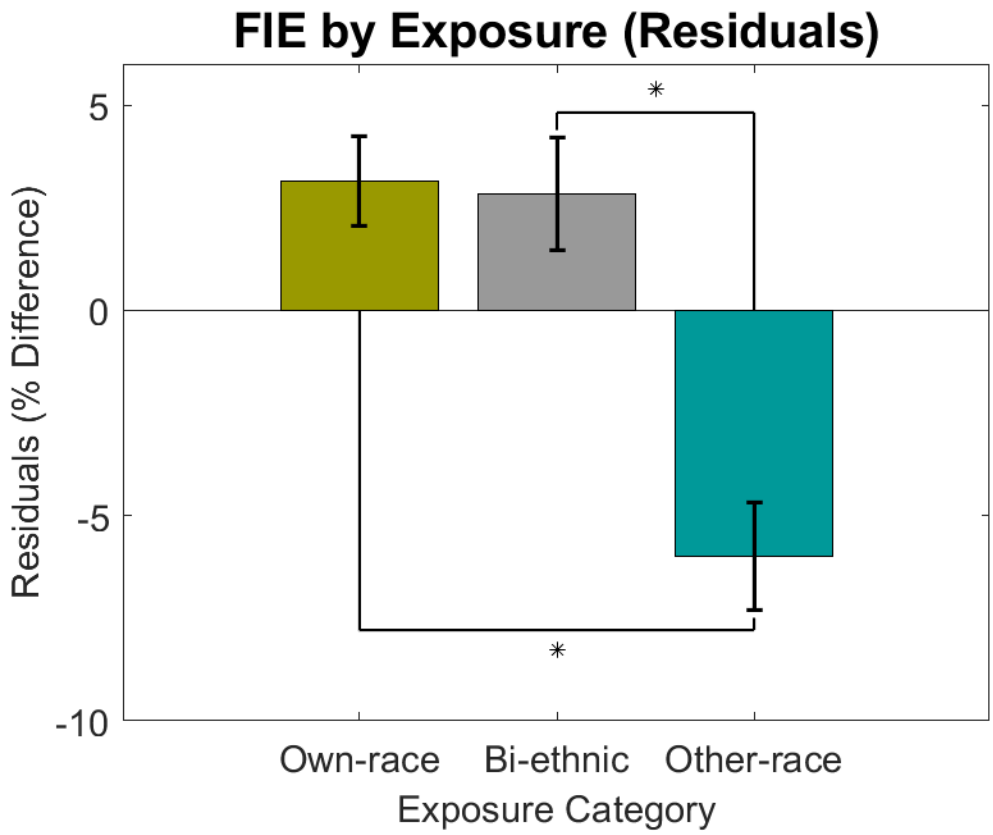

Figure 3.10 Face Inversion Effect (Residuals): Own-race vs Other-race vs Bi-ethnic Analysis. Average residual scores for own-race (left), bi-ethnic (middle), and other-race (right) exposure. Error bars represent standard errors (+/- 1SE).

3.3 Experiment 3: Identity Discrimination Task

3.3.1 General Analysis

In order to assess participants' performance based on stimulus ethnicity, we collapsed data from the female and male versions from the same stimulus ethnicity (Figure 3.11) and submitted discrimination thresholds to a repeated-measure ANOVA with stimulus category (East Asian and Caucasian) as the within-subjects factor and participant ethnicity

(Caucasian, Bi-ethnic, and East Asian) as the between-subjects factor. We observed a significant main effect of stimulus category (F (1,57) = 5.313, p = .025, η_p^2 = .085) and a significant interaction between stimulus category and participant ethnicity (F (2,57) = 6.891, p = .002, η_p^2 = .195). No main effect for participant ethnicity was detected (F (2,57) = 2.6, p = .083). Posthoc comparison of the effects of stimulus category showed that Caucasian stimuli (M = 16.02%) were significantly more difficult to discriminate (Tukey-Kramer, p = .027, d = 0.34) than East Asian stimuli (M = 14.51%). There was numerical evidence of ORE detected between the Mono-ethnic participants' performances, but this effect did not reach significant level. East Asian participants' threshold for East Asian faces (M = 14.16%) were numerically lower than Caucasian participants' threshold for the same faces (M = 17.28%) but this difference did not reach significance (Tukey-Kramer, p = .09, d = 0.70). Similarly, Caucasian participants' threshold for Caucasian faces (M = 15.38%) was numerically lower than East Asian participants' threshold (M = 17.0%) for the same faces but this difference did not reach significance (Tukey-Kramer, p = .73, d = 0.40). Interestingly, for the East Asian task, Bi-ethnic participants performed significantly better than Caucasian participants and their thresholds showed a numerical advantage over East Asian participants. Bi-ethnic group's threshold was significantly lower (M = 12.10%) for East Asian faces than Caucasian participants (Tukey-Kramer, p = .0000, d = 1.13). Moreover, Bi-ethnic participants' threshold was numerically lower than East Asian participants' threshold for the East Asian tasks but this effect was not significant (Tukey-Kramer, p = .49, d = 0.63).

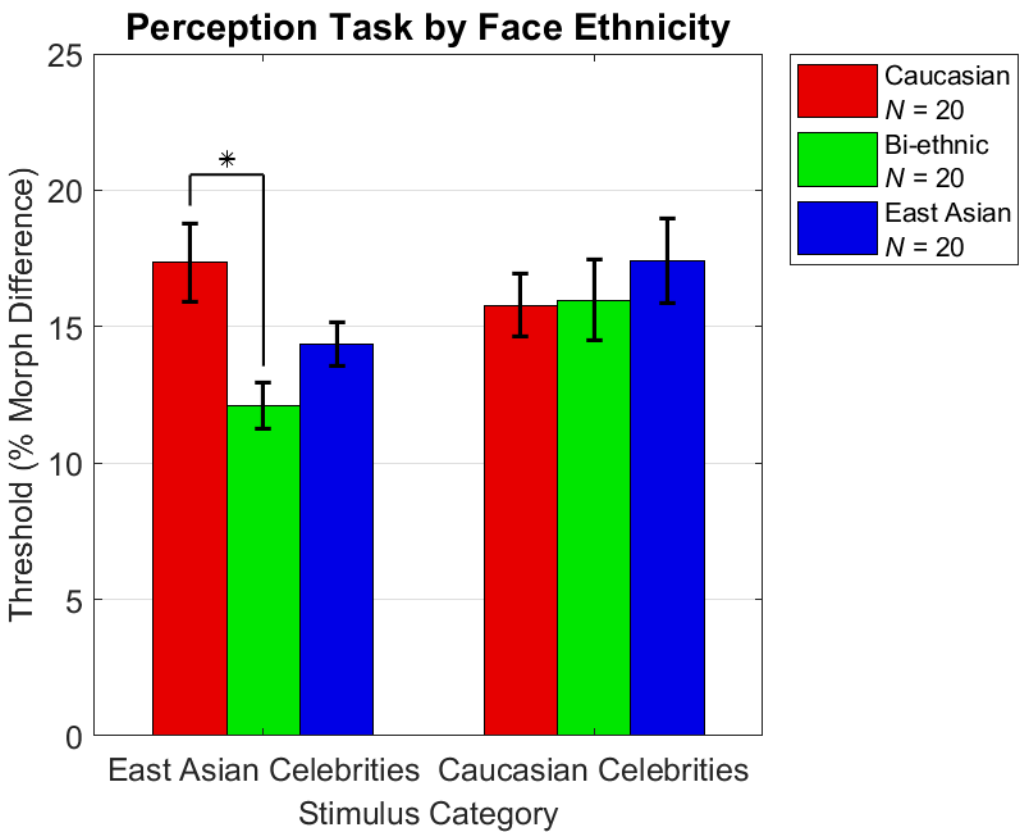

Figure 3.11 Discrimination Thresholds by Stimulus Ethnicity: General Analysis. Thresholds from male and female stimuli were averaged for each stimulus ethnicity. Error bars represent +/- 1SE.

3.3.2 Own-race vs Other-race vs Bi-ethnic Analysis

Figure 3.12 shows combined discrimination thresholds for own-race (Caucasian participants for Caucasian faces and East Asian participants for East Asian faces), other-race (Caucasian participants for East Asian faces and East Asian participants for Caucasian faces), and Bi-ethic exposure. Scores were submitted to a one-way ANOVA with exposure type (own-race, other-race, bi-ethnic) as the between-subject factor. There was a significant main effect of exposure type ($F(2,117)= 5.868$, $p = .004$, $\eta_p^2 = .091$). Further analysis using two-tailed 2 sample t-test showed significant differences between own-race and other-race exposure ($t(78) = 2.505$, $p = .0143$, $d = 0.56$) as well as bi-ethnic and other-race exposure

($t(78) = 3.036$, $p = .00326$, $d = 0.68$) and no significant difference between own-race and bi-ethnic exposure ($t(78) = 0.952$, $p = .349$).

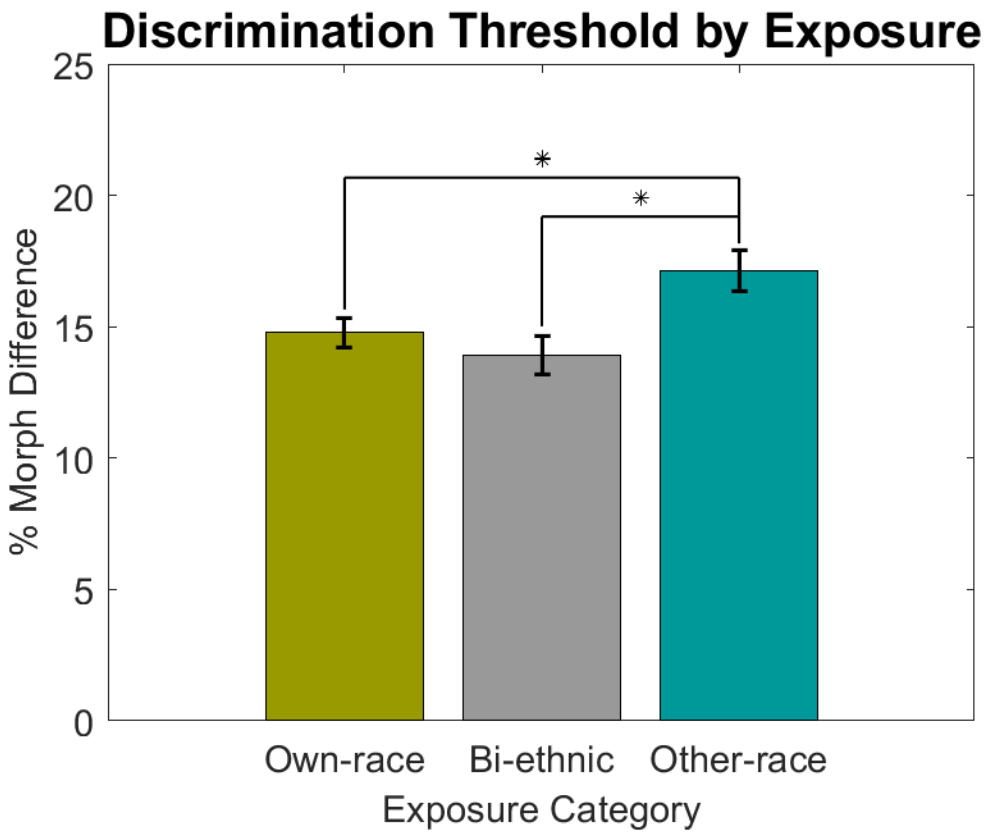

Figure 3.12 Discrimination Thresholds: Own-race vs Other-race vs Bi-ethnic Analysis. Average thresholds by stimulus ethnicity for own-race (left), bi-ethnic (middle), and other-race (right) exposure. Error bars represent standard errors (+/- 1SE).

4 Discussion

The other-race effect is a robust effect that emerges as a bias in recognition and perception abilities towards own-race faces (Meissner & Brigham, 2001; Walker & Tanaka, 2003). It is believed that this effect is to some extent due to differential contact with own-race and other-race faces. Studies have shown that this effect influences different aspects of face recognition such as face memory and perceptual abilities. The ORE is also associated with higher face inversion effect (Gilliain Rhodes et al., 2006). The inversion effect is believed to be a hallmark of expertise in visual recognition and since we are experts with upright own-race faces, the inversion effect for those faces is disproportionately large. The development of the ORE follows a phenomenon known as perceptual narrowing. At birth, newborns attend to face-like structures but do not discriminate between own- and other-race faces (Cassia et al., 2004). However, this broad perceptual ability narrows down after 3 months of age and gradually the face recognition system reaches adult expert-level performance (Kelly et al., 2007). Perceptual narrowing predicts the impacts of lack of exposure in development of visual system, however, what remains unclear is how the system will develop if it is sustainably exposed to more than one type of face ethnicity. To answer this question, we postulate three competing hypotheses. The *Resource Bottleneck* hypothesis states that the purpose behind perceptual narrowing is to avoid the split in resources and development of a generalist face recognition system, therefore all resources are focused on one type of stimuli: own-race faces which constitute the most salient input for the face recognition system. If this holds true, a divide in resources to more than one type of face-ethnicity will lead to poor performance in face recognition tasks. The *Use / Disuse* hypothesis states that perceptual narrowing happens simply because individuals have not had enough contact with other-race faces. Therefore, sustained high exposure to more than one face-ethnicity will not hinder the face recognition from reaching expert-level

performance, but will help the system to perform at a native-level for faces from different face-ethnicity. Similar to the *Use / Disuse* hypothesis, the *Facilitation* hypothesis states that perceptual narrowing happens merely due to lack of exposure. However, this account predicts that exposure to more than one face ethnicity will support richer face representations and this will lead to advantages in one or more face recognition tasks.

The current study was designed to discriminate between these alternative hypotheses. We compared Bi-ethnic participants', individuals with sustained high exposure to both East Asian and Caucasian faces, abilities across three different experiments to Caucasian and East Asian, Mono-ethnic, participants. Knowing that face memory is better for own-race faces, we examined face memory in Bi-ethnic participants for East Asian and Caucasian faces using the old/new test and the Cambridge Memory Tests, and compared their performance with Mono-ethnics. Then we examined the FIE for Caucasian and East Asian faces in Bi-ethnic participants and compared their results to own-race inversion effects in Mono-ethnics. Lastly, using an identity discrimination test, we investigated perceptual abilities of Bi-ethnic participants for Caucasian and East Asian faces and compared their performance to Mono-ethnic participants.

4.1 Experiment 1: Old/New Task

In Experiment 1 we assessed participants' face memory for Caucasian and East Asian faces using the old/new task. In congruence with previous studies (see Meissner & Brigham, 2001) our results revealed an ORE where both Caucasian and East Asian participants outperformed the other-race group (East Asian and Caucasian participants, respectively) in memorizing and recognizing own-race faces. Hence, our task and manipulation was able to produce the classical ORE effect, as expected. Importantly, Bi-ethnic performance in both the East Asian and Caucasian face categories was at near-

native level and did not significantly differ from the own-race performance of the two Mono-ethnic groups. These results do not support the Resource Bottleneck hypothesis since we found no evidence of any detrimental impact on face memory of Bi-ethnic observers on either face ethnicity. Further, the pattern of results do not provide any evidence for the *Facilitation* hypothesis, as Bi-ethnics did not demonstrate any advantage over the two Mono-ethnic groups. Finally, our finding that Bi-ethnics can achieve native-like performance on both face ethnicities is consistent with the *Use / Disuse* hypothesis.

4.2 Experiment 2: Cambridge Memory Tests

4.2.1 Upright Accuracies

In Experiment 2, we assessed participants' face recognition abilities for Caucasian and East Asian faces as well as cars using the Cambridge Memory Tests. The results from the upright Cambridge tests showed an ORE for the Mono-ethnic participants, where each of the Mono-ethnic groups had significantly higher accuracies in their own-race task compared to the other Mono-ethnic group. These results are in line with previous cross-race studies showing similar ORE in face recognition and memory using the Cambridge Face Memory paradigm (Horry, Cheong, & Brewer, 2015; Wan, Crookes, Reynolds, Irons, & McKone, 2015). Importantly, Bi-ethnic participants performed at near-native levels in both the Caucasian and East Asian tasks with accuracies statistically indistinguishable from own-race performance of both mono-ethnic groups, and significantly better than East Asian participants in the Caucasian task, near-significantly better than Caucasian participants in the East Asian task. Moreover, Bi-ethnic individuals' accuracies were not significantly different from own-race accuracies, but where significantly better than other-race accuracies. These results align closely with those of Experiment 1 and provide converging evidence that no deleterious effects follow Bi-ethnic face exposure, further diminishing the

Resource Bottleneck account. Similarly, Experiment 2 provided no evidence for the Facilitation hypothesis as no advantages in face or car memory was observed in the Bi-ethnic group. However, results from Experiment 2 show that Bi-ethnics face recognition abilities are similar to own-race and significantly better than other-race performance of Mono-ethnic participants and lend further support to the *Use / Disuse* hypothesis.

4.2.2 Inverted Accuracies and the FIE

Furthermore, we measured inversion effects across the three Cambridge tests. The subtraction-based FIE scores were significantly higher for faces than cars. This result is consistent with previous studies showing a disproportionate inversion effect for faces compared to other objects (Farah, Tanaka, & Drain, 1995; Yin, 1969). Subtractions scores showed numerical evidence for higher FIE for own-race faces but this trend did not reach significance levels. Moreover, Bi-ethnic participants' subtraction scores were not significantly different from Mono-ethnic participants' scores for own-race, hinting towards a similar FIE pattern for own-race in Bi-ethnics. In addition, we also computed a regression-based FIE score as a second index of inversion effect. An ORE was revealed in this second analysis, where both East Asian and Caucasian participants showed significantly higher inversion effects for own-race faces compared to the other Mono-ethnic group. Bi-ethnic participants' inversion effects were at near-native levels and significantly higher than those for both East Asian and Caucasian participants for other-race faces. Moreover, the combined analysis for own-race, other-race and bi-ethnic exposure revealed that Bi-ethnic participants' FIE scores both calculated from the subtraction method and method of regression were significantly better than other-race scores and were not significantly different from own-race scores. The analysis of the FIE scores lend further support for the

Use / Disuse account as no advantages or disadvantages were observed in the inversion effects of Bi-ethnics compared to the two Mono-ethnic groups.

Configural coding is considered to be a hallmark of expertise in face recognition and the FIE has been used as a measure for configural coding (Bartlett, Searcy, & Abdi, 2006; Rhodes et al., 2006). Our results show larger inversion effects for own-race faces similar to a number of previous studies (Gillian Rhodes, Brake, Taylor, & Tan, 1989; Hancock & Rhodes, 2008). Furthermore, Hancock and Rhodes (2008) found an association between higher levels of self-reported contact with other-race faces and smaller configural coding differences. Our analysis of regression-based FIE scores show a similar pattern to this finding, as Bi-ethnic participants show higher inversion effects for other-race faces compared to Mono-ethnics.

4.3 Experiment 3: Identity Discrimination Task

In Experiment 3, we examined participants' ability for discriminating between pairs of similar familiar faces of East Asian and Caucasian celebrities. For each pair of celebrity identities, discrimination thresholds were measured around the individual point of subjective equality for each participant. The results from the Caucasian celebrity pairs did not reveal any ORE as no significant differences were detected across the three participant groups. However, the results from the East Asian celebrity pairs did produce a significant ORE with both East Asian and Caucasian participants showing lower discrimination thresholds for own-race faces compared to the other mono-ethnic group. In addition, Bi-ethnic participants performed at native levels for all tasks, and significantly better than Caucasian participants for the East Asian tasks. Interestingly, Bi-ethnic participants performed numerically, but not significantly, better than East-Asian participants in the East Asian tasks. Moreover, the combined analysis for own-race, other-race and bi-ethnic exposures showed an ORE where

thresholds for own-race performance were significantly lower than other-race. This analysis also revealed that Bi-ethnic participants' thresholds were significantly lower than other-race thresholds and not significantly different from own-race thresholds. These pattern of results in face perception lend further support the *Use / Disuse* hypothesis, and perhaps hint at the possibility of the *Facilitation* hypothesis.

Walker and Tanaka (2003) in their study of face perception abilities across East Asian and Caucasian participants, used a morph continuum of Caucasian and East Asian unfamiliar faces to conduct a same / difference discrimination task. Their results showed that each participant group performed better at discriminating own-race faces than for other-race faces but no comparisons were made between the two participant groups. Using a similar approach Walker and Hewstone (2008) showed that an ORE for South Asian faces was present in Caucasian participants but a similar ORE for Caucasian faces was not present in South Asian participants. They argued that the South Asian participants had more contact with other-race faces and therefore have discriminated both face types with no differences. In our study we used different morph levels from faces with the same ethnicity and compared performances across different participant groups for the same task rather than comparing the same participant group's performance across different tasks. However, our experiment failed to produce an ORE for the Caucasian task. This can be, in part, attributed to the differences in media exposures between our Caucasian and East Asian participants, as well as differences in the visibility and celebrity status of the Caucasian faces vs. the East Asian faces. While our Caucasian participants spent significantly more weekly hours on programs with Caucasian casts than programs with East Asian casts (two-tailed paired t-test, $p = 0.0001$), East Asian participants did not show any significant differences between the time they spent on programs with Caucasian and East Asian cast

(two-tailed paired t-test, $p = 0.7124$). This equal exposure to East Asian and Caucasian media within the East Asian group could have resulted in a better discrimination performance for other-race celebrities. Another explanation for these results could the difference in the level of fame among the Caucasian and East Asian celebrities. Although we made sure that the participants are familiar with all celebrity identities prior to running the tests, Caucasian celebrities' presence in movies and media is more widespread than East Asian celebrities, hence potentially leading biases in exposure. Nevertheless, the significant difference observed through this study was Bi-ethnic participants' advantage in the East Asian task over Caucasian participants.

5 Conclusions

Our cross-race comparison of face memory using the old/new task, the Cambridge memory task, the FIE, and the perceptual identity discrimination task, all demonstrate that Bi-ethnic participants perform at a native level for faces of both races on par with own-race performance of Mono-ethnic participants and bi-ethnic exposure does not have a detrimental impact on face recognition abilities based on ethnicity. In light of our competing hypotheses, these results provide converging evidence supporting the Use/Disuse hypothesis. None of the measures in our study, which encompass measures of face memory, configural coding, and face perception, showed any indication of detrimental effects on face (or car) abilities of Bi-ethnic observers. Based on this converging evidence, we argue that sustained high exposure to two face-ethnicities does not impose any limitations on face recognition abilities and conclusively rule out the Resource Bottleneck hypothesis, at least for the case of two ethnicities. Future studies assessing effects of sustained exposure to three or more face ethnicities can further examine this conclusion.

All but one of our measures showed that Bi-ethnics performed at native or near-native levels. In two of the four perceptual discrimination tasks, the female and male East Asian tasks, Bi-ethnics not only outperformed Caucasians, but they also showed enhanced performance compared to the own-race performance of East Asian participants, though this latter effect did not reach significance. This result hints at potential advantages of sustained bi-ethnic exposure in face perception and provides partial support for the *Facilitation* hypothesis. Future studies can examine other perceptual tasks, including identity discrimination of unfamiliar faces, to further explore the potential of the Facilitation hypothesis.

References

Anastasi, J. S., & Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic bulletin & review*, *12*(6), 1043-1047.

Balas, B., & Saville, A. (2015). N170 face specificity and face memory depend on hometown size. Neuropsychologia, 69, 211-217.

Bartlett, J. C., Searcy, J. H., & Abdi, H. (2003). What are the routes to face recognition. *Perception of faces, objects and scenes: Analytic and holistic processes*, 21-52.

Behrmann, M., Winocur, G., & Moscovitch, M. (1992). Dissociation between mental imagery and object recognition in a brain-damaged patient. *Nature*, *359*(6396), 636.

Bothwell, R. K., Brigham, J. C., & Malpass, R. S. (1989). Cross-racial identification. *Personality and Social Psychology Bulletin*, *15*(1), 19-25.

Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial vision*, *10*, 433-436.

Bruce, V., Doyle, T., Dench, N., & Burton, M. (1991). Remembering facial configurations. *Cognition*, *38*(2), 109-144.

Cassia, V. M., Turati, C., & Simion, F. (2004). Can a nonspecific bias toward top-heavy patterns explain newborns' face preference?. Psychological Science, 15(6), 379-383.

Chiroro, P., & Valentine, T. (1995). An investigation of the contact hypothesis of the own-race bias in face recognition. The Quarterly Journal of Experimental Psychology Section A, 48(4), 879-894.

Damasio, A. R., Tranel, D., & Damasio, H. (1990). Face agnosia and the neural substrates of memory. Annual review of neuroscience, 13(1), 89-109.

DeGutis, J., Wilmer, J., Mercado, R. J., & Cohan, S. (2013). Using regression to measure holistic face processing reveals a strong link with face recognition ability. Cognition, 126(1), 87-100.

Dennett, H. W., McKone, E., Tavashmi, R., Hall, A., Pidcock, M., Edwards, M., & Duchaine, B. (2012). The Cambridge Car Memory Test: A task matched in format to the Cambridge Face Memory Test, with norms, reliability, sex differences, dissociations from face memory, and expertise effects. Behavior Research Methods, 44(2), 587-605.

Desimone, R. (1991). Face-selective cells in the temporal cortex of monkeys. Journal of cognitive neuroscience, 3(1), 1-8.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of expertise. Journal of Experimental Psychology: General, 115(2), 107.

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. Neuropsychologia, 44(4), 576-585.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. Science, 171(3968), 303-306.

Elliott, E. S., Wills, E. J., & Goldstein, A. G. (1973). The effects of discrimination training on the recognition of white and oriental faces. Bulletin of the Psychonomic Society, 2(2), 71-73.

Evans, J. R., Marcon, J. L., & Meissner, C. A. (2009). Cross-racial lineup identification: Assessing the potential benefits of context reinstatement. Psychology, Crime & Law 15(1), 19-28.

Fair, J., Flom, R., Jones, J., & Martin, J. (2012). Perceptual learning: 12-month-olds' discrimination of monkey faces. Child development, 83(6), 1996-2006.

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

Farroni, T., Johnson, M. H., Menon, E., Zulian, L., Faraguna, D., & Csibra, G. (2005). Newborns' preference for face-relevant stimuli: Effects of contrast polarity. Proceedings of the National Academy of Sciences, 102(47), 17245-17250.

Feinman, S., & Entwisle, D. R. (1976). Children's ability to recognize other children's faces. Child development, 506-510.

Ferguson, K. T., Kulkofsky, S., Cashon, C. H., & Casasola, M. (2009). The development of specialized processing of own-race faces in infancy. Infancy, 14(3), 263-284.

Furl, N., Phillips, P. J., & O'Toole, A. J. (2002). Face recognition algorithms and the other-race effect: computational mechanisms for a developmental contact hypothesis. Cognitive Science, 26(6), 797-815.

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(8), 845.

Goldstein, A. G. (1965). Learning of inverted and normally oriented faces in children and adults. Psychonomic Science, 3(1-12), 447-448.

Goldstein, A. G., & Chance, J. E. (1985). Effects of training on Japanese face recognition: Reduction of the other-race effect. Bulletin of the Psychonomic Society, 23(3), 211-214.

Gottlieb, G. (1976). The roles of experience in the development of behavior and the nervous system. In Studies on the development of behavior and the nervous system (Vol. 3, pp. 25-54). Elsevier.

Hainline, L. (1978). Developmental changes in visual scanning of face and nonface patterns by infants. Journal of Experimental Child Psychology, 25(1), 90-115.

Haith, M. M., Bergman, T., & Moore, M. J. (1977). Eye contact and face scanning in early infancy. Science, 198(4319), 853-855.

Hancock, K. J., & Rhodes, G. (2008). Contact, configural coding and the other-race effect in face recognition. British Journal of Psychology, 99(1), 45-56.

Hayden, A., Bhatt, R. S., Reed, A., Corbly, C. R., & Joseph, J. E. (2007). The development of expert face processing: Are infants sensitive to normal differences in second-order relational information?. Journal of Experimental Child Psychology, 97(2), 85-98.

Hochberg, J., & Galper, R. E. (1967). Recognition of faces: I. An exploratory study. Psychonomic Science, 9(12), 619-620.

Horry, R., Cheong, W., & Brewer, N. (2015). The other-race effect in perception and recognition: insights from the complete composite task. Journal of Experimental Psychology: Human Perception and Performance, 41(2), 508.

Jackiw, L. B., Arbuthnott, K. D., Pfeifer, J. E., Marcon, J. L., & Meissner, C. A. (2008). Examining the cross-race effect in lineup identification using Caucasian and First Nations samples. Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 40(1), 52.

Jaquet, E., Rhodes, G., & Hayward, W. G. (2008). Race-contingent aftereffects suggest distinct perceptual norms for different race faces. Visual Cognition, 16(6), 734-753.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. Journal of neuroscience, 17(11), 4302-4311.

Kelly, D. J., Liu, S., Lee, K., Quinn, P. C., Pascalis, O., Slater, A. M., & Ge, L. (2009). Development of the other-race effect during infancy: Evidence toward universality?. Journal of experimental child psychology, 104(1), 105-114.

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.

Kleiner, K. A. (1987). Amplitude and phase spectra as indices of infants' pattern preferences. Infant Behavior and Development, 10(1), 49-59.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., & Murray, R. (2007). What's new in Psychtoolbox-3? Perception. In Thirtieth European Conference on Visual Perception, Arezzo, Italy, August.

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(4), 811.

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. Science, 218(4577), 1138-1141.

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), 14-14.

Lamont, A. C., Stewart-Williams, S., & Podd, J. (2005). Face recognition and aging: Effects of target age and memory load. Memory & Cognition, 33(6), 1017-1024.

Lewkowicz, D. J., & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. Trends in cognitive sciences, 13(11), 470-478.

Lindsay, D. S., Jack, P. C., & Christian, M. A. (1991). Other-race face perception. Journal of applied psychology, 76(4), 587.

Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. Journal of personality and social psychology, 13(4), 330.

Maurer, D., & Werker, J. F. (2014). Perceptual narrowing during infancy: A comparison of language and faces. Developmental Psychobiology, 56(2), 154-178.

Maurer, D., & Salapatek, P. (1976). Developmental changes in the scanning of faces by young infants. Child development, 523-527.

McKone, E., Stokes, S., Liu, J., Cohan, S., Fiorentini, C., Pidcock, M., ... & Pelleg, M. (2012). A robust method of measuring other-race and other-ethnicity effects: The Cambridge Face Memory Test format. PLoS One, 7(10), e47956.

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.

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

Mondloch, C. J., Lewis, T. L., Budreau, D. R., Maurer, D., Dannemiller, J. L., Stephens, B. R., & Kleiner-Gathercoal, K. A. (1999). Face perception during early infancy. Psychological Science, 10(5), 419-422.

Oruc, I., Shafai, F., & Iarocci, G. (2018). Link Between Facial Identity and Expression Abilities Suggestive of Origins of Face Impairments in Autism: Support for the Social-Motivation Hypothesis. Psychological science, 29(11), 1859-1867.

Oruc, I., Shafai, F., Murthy, S., Lages, P., & Ton, T. (2018). The adult face-diet: A naturalistic observation study. Vision research.

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.

Pascalis, O., Scott, L. S., Kelly, D. J., Shannon, R. W., Nicholson, E., Coleman, M., & Nelson, C. A. (2005). Plasticity of face processing in infancy. Proceedings of the national academy of sciences, 102(14), 5297-5300.

Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life?. Science, 296(5571), 1321-1323.

Patterson, M. L., & Werker, J. F. (2003). Two-month-old infants match phonetic information in lips and voice. Developmental Science, 6(2), 191-196.

Pelli, D. G., & Vision, S. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spatial vision, 10, 437-442.

Perrett, David Ian, Jari K. Hietanen, Michael W. Oram, and Philip J. Benson. "Organization and functions of cells responsive to faces in the temporal cortex." Philosophical transactions of the royal society of London. Series B: Biological sciences 335, no. 1273 (1992): 23-30.

Phillips, P. J., Jiang, F., Narvekar, A., Ayyad, J., & O'Toole, A. J. (2011). An other-race effect for face recognition algorithms. ACM Transactions on Applied Perception (TAP), 8(2), 14.

Pons, F., Lewkowicz, D. J., Soto-Faraco, S., & Sebastián-Gallés, N. (2009). Narrowing of intersensory speech perception in infancy. Proceedings of the National Academy of Sciences, 106(26), 10598-10602.

Regan, D., & Beverley, K. I. (1985). Postadaptation orientation discrimination. JOSA A, 2(2), 147-155.

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.

Rhodes, G., Brake, S., Taylor, K., & Tan, S. (1989). Expertise and configural coding in face recognition. British journal of psychology, 80(3), 313-331.

Robotham, R. J., & Starrfelt, R. (2018). Tests of whole upright face processing in prosopagnosia: A literature review. Neuropsychologia.

Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces?. Behavioral and cognitive neuroscience reviews, 1(1), 63-75.

Sangrigoli, S., Pallier, C., Argenti, A. M., Ventureyra, V. A. G., & de Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. Psychological Science, 16(6), 440-444.

Sangrigoli, S., & De Schonen, S. (2004a). Recognition of own-race and other-race faces by three-month-old infants. Journal of Child Psychology and Psychiatry, 45(7), 1219-1227.

Sangrigoli, S., & De Schonen, S. (2004b). Effect of visual experience on face processing: A developmental study of inversion and non-native effects. Developmental Science, 7(1), 74-87.

Scapinello, K. F., & Yarmey, A. D. (1970). The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. Psychonomic Science, 21(6), 329-330.

Schultz, R. T. (2005). Developmental deficits in social perception in autism: the role of the amygdala and fusiform face area. International Journal of Developmental Neuroscience, 23(2-3), 125-141.

Scott, L. S., & Monesson, A. (2009). The origin of biases in face perception. Psychological Science, 20(6), 676-680.

Shafai, F., & Oruc, I. (2018). Qualitatively similar processing for own-and other-race faces: Evidence from efficiency and equivalent input noise. Vision research, 143, 58-65.

Simpson, E. A., Varga, K., Frick, J. E., & Fragaszy, D. (2011). Infants experience perceptual narrowing for nonprimate faces. Infancy, 16(3), 318-328.

Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories. Psychology, Public Policy, and Law, 7(1), 36.

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.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. The Quarterly journal of experimental psychology, 46(2), 225-245.

Tanaka, J. W., Heptonstall, B., & Hagen, S. (2013). Perceptual expertise and the plasticity of other-race face recognition. Visual Cognition, 21(9-10), 1183-1201.

Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. Cognition, 93(1), B1-B9.

Tanaka, J. W., & Pierce, L. J. (2009). The neural plasticity of other-race face recognition. Cognitive, Affective, & Behavioral Neuroscience, 9(1), 122-131.

Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. Journal of cognitive neuroscience, 3(1), 71-86.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. British journal of psychology, 79(4), 471-491.

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

Vouloumanos, A., Hauser, M. D., Werker, J. F., & Martin, A. (2010). The tuning of human neonates' preference for speech. Child development, 81(2), 517-527.

Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. Developmental science, 10(2), 159-164.

Walker, P. M., & Hewstone, M. (2006). A perceptual discrimination investigation of the own-race effect and intergroup experience. Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 20(4), 461-475.

Walker, P. M., & Tanaka, J. W. (2003). An encoding advantage for own-race versus other-race faces. Perception, 32(9), 1117-1125.

Wan, L., Crookes, K., Reynolds, K. J., Irons, J. L., & McKone, E. (2015). A cultural setting where the other-race effect on face recognition has no social–motivational component and derives entirely from lifetime perceptual experience. Cognition, 144, 91-115.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. Perception & psychophysics, 33(2), 113-120.

Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. Child development, 349-355.

Wiese, H., Kaufmann, J. M., & Schweinberger, S. R. (2012). The neural signature of the own-race bias: Evidence from event-related potentials. Cerebral Cortex, 24(3), 826-835.

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: The Official Journal of the Society for Applied Research in Memory and Cognition, 17(3), 365-373.

Yin, R. K. (1969). Looking at upside-down faces. Journal of experimental psychology, 81(1), 141.

Young, A. W., Hellawell, D., & Hay, D. C. (2013). Configurational information in face perception. Perception, 42(11), 1166-1178.

Yurovsky, D., Fricker, D. C., Yu, C., & Smith, L. B. (2014). The role of partial knowledge in statistical word learning. Psychonomic bulletin & review, 21(1), 1-22.