FACE PROCESSING IN MALAYSIAN CHINESE ADULTS

Chrystalle Bih Yuan Tan

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

June 2014

Abstract

Cross-cultural studies have identified a distinct holistic-analytic pattern that observers employ in various cognitive and perceptual tasks. Recent face perception studies utilizing eye tracking methodologies have also revealed distinct Eastern and Western viewing patterns when recognizing identities and emotions. However, studies exploring genetic and cultural factors found that British born Chinese observers employed either Eastern or Western eye movement strategies, suggesting that a simple Eastern-Western distinction does not fully explain the diversity in observers' eye movement strategies. Although Malaysia is an East Asian country, it is strongly multicultural and heavily influenced by Western culture. This thesis aimed to investigate Malaysian Chinese participants' eye movement strategy and recognition abilities by requiring participants to perform static and dynamic face recognition (Chapter Three and Chapter Five) and emotion recognition (Chapter Four and Chapter Six) tasks on African, East Asian, and Western Caucasian faces.

Possibly due to greater exposure to Western culture, Malaysian Chinese adopted a combination of Eastern and Western eye movement strategies to better recognize static identities and emotions, indicating that individuals learn to use eye movement strategies that are optimized for recognizing the faces with which they are more familiar. Dynamic face processing tasks revealed different fixation patterns, suggesting that the nose may be a more optimal strategy to recognize identities whereas the mouth provides important cues for identifying emotions. These findings imply that directing attention towards facial cues that have the highest diagnostic value in relation to a specific race of face and/or

task may reduce the other-race effect and improve recognition performance. Additionally, the type of stimuli (static versus dynamic) and nature of task (face recognition versus emotion recognition) rendered different eye movement strategies from participants, providing further support that participants perform differently when presented with dynamic stimuli that are more naturalistic and when instructed to perform different tasks.

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Dr. Elizabeth Sheppard and Dr. Ian Stephen for their supervision, guidance, and support throughout my Ph.D. I would also like extend my appreciation to other members of staff for their invaluable advice and guidance; thanks to Dr. Jessica Price for taking her time to read and critique my thesis and Dr. Kirsten McKenzie for her advice on data analysis. Thanks to the Psychology Department in Sunway University for providing me a laboratory to collect data on their campus and to all of my participants who took their time to take part in the project. Special thanks to my friends and fellow PhD students for their continual support and encouragement. And finally, I would like to thank my family for their unconditional love and support; they have been my continual source of inspiration and motivation.

Table of Contents

Abstract ii

Acknowledgements iv

Chapter One: Face Processing 1

1.1. Main Theoretical Approaches to Face Recognition 2

1.1.1. Feature Detection Theory 2

1.1.2. Configural Face Processing Theory 6

1.1.2.1. Face Inversion Effect 7

1.1.2.2. Composite Face Effect 10

1.1.2.3. Part-Whole Effect 12

1.1.3. Information Processing Models 14

1.2. The Theoretical Framework of Facial Expressions of Emotion 20

1.2.1. The Universality Hypothesis 21

1.2.2. The Culture Specificity Hypothesis 22

1.2.3. Ekman's Theoretical Framework 24

1.3. Ecological Validity 26

1.3.1. Eyewitness Testimony 26

1.3.2. Methodological Considerations 29

1.4. Summary 33

Chapter Two: Cultural Differences 36

2.1. The Holistic and Analytic Dichotomy 36

2.1.1. Cultural Differences in Cognition 37

2.1.2. Cultural Differences in Visual Perception 39

2.1.3. Summary 42

2.2. Cultural Differences in Face Perception 43
2.3. The Malaysian Culture 47
2.4. Aims of This Thesis 49
2.4.1. Methodological Considerations 50
Chapter Three: Static Face Recognition 53
3.1. Introduction 53
3.1.1. Other-Race Effect 54
3.1.2. Eye Tracking 60
3.2. Current Study 63
3.3. Methods 64
3.3.1. Participants 64
3.3.2. Materials 65
3.3.3. Procedure 66
3.3.4. Eye Tracking 67
3.3.5. Data Analyses 67
3.4. Results 70
3.4.1. Recognition Sensitivity 70
3.4.2. Eye Tracking Results 72
3.4.2.1. Total Number of First Fixations 74
3.4.2.2. Total Fixation Count 76
3.4.2.3. Mean Fixation Duration 78
3.5. Discussion 80
Chapter Four: Static Emotion Recognition 86
4.1. Introduction 87

4.1.1. Theoretical Framework 87
4.1.2. Identification of Important Facial Cues Used to Categorize Emotions 89
4.2. Current Study 94
4.3. Methods 95
4.3.1. Participants 95
4.3.2. Materials 96
4.3.3. Procedure 98
4.3.4. Eye Tracking 99
4.3.5. Data Analyses 99
4.4. Results 101
4.4.1. Categorization Sensitivity 101
4.4.2. Eye Tracking Results 103
4.4.2.1. Total Number of First Fixations 106
4.4.2.2. Total Fixation Count 109
4.4.2.3. Mean Fixation Duration 112
4.5. Static Face Recognition and Emotion Recognition Tasks 116
4.5.1. Total Number of First Fixations 119
4.5.2. Total Fixation Count 121
4.5.4. Mean Fixation Duration 123
4.6. Discussion 126
Chapter Five: Dynamic Face Recognition 135
5.1. Introduction 136
5.1.1. Facial Motions and Social Signals 136

5.1.1.1. The Representation Enhancement and Supplemental Information Hypotheses 137
5.1.2. Eye Tracking 140
5.2. Current Study 143
5.3. Methods 144
5.3.1. Participants 144
5.3.2. Materials 144
5.3.3. Procedure 146
5.3.4. Eye Tracking 146
5.3.5. Data Analyses 147
5.4. Results 147
5.4.1. Recognition Sensitivity 147
5.4.2. Eye Tracking Results 149
5.4.2.1. Total Number of First Fixations 151
5.4.2.2. Total Fixation Count 152
5.4.2.3. Mean Fixation Duration 153
5.5. Discussion 154
Chapter Six: Dynamic Emotion Recognition 159
6.1. Introduction 160
6.1.1. Dynamic Advantage Hypothesis 160
6.1.1.1. Supporting Evidence 161
6.1.1.2. Conflicting Evidence 162
6.1.2. Differences in the Portrayal and Processing of Facial Expressions 165
6.2. Current Study 167

6.3. Methods 169
6.3.1. Participants 169
6.3.2. Materials 169
6.3.3. Procedure 173
6.3.4. Eye Tracking 174
6.3.5. Data Analyses 174
6.4. Results 175
6.4.1. Categorization Sensitivity 175
6.4.2. Eye Tracking Results 179
6.4.2.1. Total Number of First Fixations 181
6.4.2.2. Total Fixation Count 184
6.5. Dynamic Face Recognition and Emotion Recognition Tasks 186
6.5.1. Total Number of First Fixations 189
6.5.2. Total Fixation Count 189
6.5.3. Mean Fixation Duration 191
6.6. Discussion 192
Chapter Seven: General Discussion and Conclusions 198
7.1. Overview of Research 198
7.2. Summary and Discussion of Results 201
7.2.1. Chapter Three: Static Face Recognition 201
7.2.2. Chapter Four: Static Emotion Recognition 202
7.2.3. Static Face Recognition and Emotion Recognition Tasks 204
7.2.4. Chapter Five: Dynamic Face Recognition 207
7.2.5. Chapter Six: Dynamic Emotion Recognition 209

7.2.6. Dynamic Face Recognition and Emotion Recognition Tasks 210

7.3. Implications of Results 212

7.4. Limitations and Suggestions for Future Research 223

7.5. Conclusion 227

References 229

Appendix A: Results for Total Fixation Duration in the Static Face Recognition Task 244

Appendix B: Results for Total Fixation Duration in the Static Emotion Recognition Task 247

Appendix C: Results for Total Fixation Duration in the Static Face Recognition and Emotion Recognition Tasks 253

Appendix D: Results for Total Fixation Duration in the Dynamic Face Recognition Task 256

Appendix E: Results for Total Fixation Duration and Mean Fixation Duration in the Dynamic Emotion Recognition Task 258

Appendix F: Results for Total Fixation Duration in the Dynamic Face Recognition and Emotion Recognition Tasks 262

List of Figures

Figure 1.1. Examples of paired Identi-Kit faces differing in two, four, and seven features in both photographically positive and negative forms used in Bradshaw and Wallace's (1971) study. 4

Figure 1.2. Examples of the 8 photo-fit faces used in Sergent's (1984) study; C = contour, E = eyes, I = internal space. 5

Figure 1.3. Examples of whole and part masks in Farah, Wilson, Drain, & Tanaka's (1998) study. 9

Figure 1.4. Examples of (aligned and misaligned) stimuli and procedure used in Turati, di Giorgio, Bardi, &Simion's (2010) study. 12

Figure 1.5. Examples of isolated parts, intact faces, and scrambled faces in Tanaka & Farah's (1993) study. 13

Figure 1.6. A functional model of face processing proposed by Bruce and Young (1986). 16

Figure 1.7. The interactive activation model of person recognition proposed by Burton, Bruce, and Johnson (1991). 17

Figure 1.8. Examples of face and name stimuli from the same person (top), related (middle), and unrelated (bottom) task conditions in De Haan, Young, &Newcombe's (1987) study. 19

Figure 3.1. Histogram of British born Chinese participants' classifier result. 63

Figure 3.2. The predefined area of interest (AOI) used to analyze eye gaze. 70

Figure 3.3. Malaysian Chinese participants' recognition sensitivity for African, East Asian, and Western Caucasian faces. 71

Figure 3.4. Average number of first fixations landing on the eyes, mouth, and nose during the learning and recognition tasks. 74

Figure 3.5. Average total fixation count for the eyes, mouth, and nose during the learning and recognition tasks. 77

Figure 3.7. Mean fixation duration on the eyes, mouth, and nose during the learning and recognition tasks, in seconds..... 79

Figure 4.1. Examples of the creation of bubblized image (a) and video (b) used Blais et al. (2012)'s study. 94

Figure 4.2. Examples of stimuli used in the experiment..... 98

Figure 4.3. The predefined area of interest (AOI) used to analyze eye gaze..... 101

Figure 4.4. Malaysian Chinese participants' categorization sensitivity for East Asian and Western Caucasian emotions. 103

Figure 4.5. Average number of first fixations landing on the eyes, mouth, and nose during the emotion recognition task..... 108

Figure 4.6. Average total fixation count for the eyes, mouth, and nose during the emotion recognition task. 111

Figure 4.7. Mean fixation duration for the eyes, mouth, and nose during the emotion recognition task, in seconds. 115

Figure 4.8. Average number of first fixations landing on the eyes, mouth, and nose during the face recognition and emotion recognition tasks. 119

Figure 4.9. Average total fixation count for the eyes, mouth, and nose during the face recognition and emotion recognition tasks. 122

Figure 4.10. Average total fixation count for the eyes, mouth, and nose during the emotion recognition and face recognition tasks..... 123

Figure 4.11. Mean fixation duration for the eyes, mouth, and nose during the face recognition and emotion recognition tasks, in seconds. 124

Figure 5.1. Examples of stimuli used in the experiment..... 145

Figure 5.2. Malaysian Chinese participants' recognition sensitivity for African, East Asian, and Western Caucasian faces. 148

Figure 5.3. Average number of first fixations landing on the eyes, mouth, and nose during the learning and recognition tasks..... 151

Figure 5.4. Average total fixation count for the eyes, mouth, and nose during the learning and recognition tasks. 153

Figure 5.5. Mean fixation duration for the eyes, mouth, and nose during the learning and recognition tasks, in seconds. 154

Figure 6.1. Examples of stimuli used in the experiment. 173

Figure 6.2. Malaysian Chinese participants' categorization sensitivity for African, East Asian, and Western Caucasian emotions. 179

Figure 6.3. Average number of first fixations landing on the eyes, mouth, and nose during the emotion recognition task. 183

Figure 6.4. Average total fixation count for the eyes, mouth, and nose during the emotion recognition task. 185

Figure 6.5. Average number of first fixations landing on the eyes, mouth, and nose during the face recognition and emotion recognition tasks. 189

Figure 6.6. Average total fixation count for the eyes, mouth, and nose during the face recognition and emotion recognition tasks. 190

Figure 6.7. Mean fixation duration for the eyes, mouth, and nose during the face recognition and emotion recognition tasks, in seconds. 191

Figure A.1. Average total fixation duration for the eyes, mouth, and nose during the learning and recognition tasks, in seconds. 245

Figure B.1. Average total fixation duration for the eyes, mouth, and nose during the emotion recognition task, in seconds. 252

Figure C.1. Average total fixation duration for the eyes, mouth, and nose during the face recognition and emotion recognition tasks, in seconds. 253

Figure D.1. Average total fixation duration for the eyes, mouth, and nose during the learning and recognition tasks, in seconds. 257

Figure E.1. Average total fixation duration for the eyes, mouth, and nose during the emotion recognition task, in seconds. 259

Figure E.2. Average mean fixation duration for the eyes, mouth, and nose during the emotion recognition task, in seconds. 261

Figure F.1. Average total fixation duration for the eyes, mouth, and nose during the face recognition and emotion recognition tasks, in seconds. 262

List of Tables

Table 3.1. Hit rates and false alarm rates in recognizing African, East Asian, and Western Caucasian faces. 71

Table 3.2. Summary of statistics for each dependent variable in the linear mixed models. 73

Table 4.1. Hit (H) rates and false alarm (FA) rates in recognizing East Asian (EA) and Western Caucasian (WC) facial expressions of “neutral,” “happy,” “surprise,” “fear,” “disgust,” “anger,” and “sad” faces. 101

Table 4.2. Summary of statistics for each dependent variable in the linear mixed models. 105

Table 4.3. Summary of statistics for each dependent variable in the linear mixed models. 118

Table 5.1. Hit rates and false alarm rates in recognizing African, East Asian, and Western Caucasian faces. 148

Table 5.2. Summary of statistics for each dependent variable in the linear mixed models. 150

Table 6.1. The 42 chosen African (A), East Asian (EA), and Western Caucasian (WC) models’ average self-rating of the intensity of emotion felt during the narrated experience and the descriptiveness of the narration. 170

Table 6.2. Raters’ average rating of the 42 chosen African (A), East Asian (EA), and Western Caucasian (WC) models’ intensity of emotion when narrating an emotional experience. 172

Table 6.3. Hit (H) rates and false alarm (FA) rates in recognizing African (A), East Asian (EA), and Western Caucasian (WC) expressions of “neutral,” “happy,” “surprise,” “fear,” “disgust,” “anger,” and “sad” faces. 177

Table 6.4. Summary of statistics for each dependent variable in the linear mixed models. 180

Table 6.5. Summary of statistics for each dependent variable in the linear mixed models. 188

Chapter One

Face Processing

Faces play an important role in our lives; they not only act as a means of identification, but also provide important information about an individual, including emotion, age, gender, ancestry, health status, and attractiveness. Although faces consist of the same basic features arranged in a similar pattern, observers can easily distinguish between faces; an ability that is not necessarily transferrable to other animals or objects (e.g. most people may not be able to distinguish one bird amongst a flock of birds or one tree in a forest). The ability to recognize faces may have social and evolutionary advantages, such as the ability to recognize cheaters or identify allies (Volstorf, Rieskamp, & Stevens, 2011; Mealey, Daood, &Krage, 1996); therefore, examining how observers process faces is informative in understanding human social behavior (Peterson & Rhodes, 2003).

This thesis investigates Malaysian Chinese participants' eye movement strategy and accuracy in recognizing static photographs and dynamic videos of own- and other-race identities and facial expressions of emotion. The current chapter will focus on the three main theoretical approaches to face recognition and the theoretical framework of facial expressions of emotion while Chapter Two will focus on the relationship between culture and face processing tasks such as face recognition and emotion recognition. In this thesis, the term "holistic" will be used to describe the context dependent cognitive style which is global and has high context sensitivity with greater focus on gist and/or configural information (e.g. the relative positions of the features on the face) whereas the term "analytic" will be used to describe the context independent cognitive style which is local and has low context sensitivity with

greater focus on detail and/or individual features (Happé & Frith, 2006; Nisbett & Miyamoto, 2005; Piepers & Robbins, 2012).

1.1. Main Theoretical Approaches to Face Recognition

In the next sections, the three main theoretical approaches to face recognition (i.e. feature detection theory, configural face processing theory, and information processing models) will be discussed in more detail.

1.1.1. Feature Detection Theory

Each person has a unique set of features that, once identified by an observer, can be used to recognize the person. The feature detection theory proposed that observers build an image of a face by analyzing individuating information provided by each feature (e.g. eyes, nose, or mouth). Some studies have provided evidence that features play a major role in face recognition and different features may affect recognition performance in different ways. For instance, a study by Davies, Ellis, and Shepherd (1977) found that certain features of the face are more important than others. Utilizing Photofit Kit to construct a series of faces, participants were shown a target face and were then required to identify it from a series of distractor faces, which differed from the target face in only one feature. Results showed that participants were most aware of changes in the eyes, followed by the mouth then nose, suggesting that the eyes are the most important features for face recognition. A previous study has also found that different features are important for familiar and unfamiliar faces; more specifically, internal features were more useful than external features for the recognition of familiar faces but both internal and external features were equally useful for the recognition of unfamiliar faces (Ellis, Shepherd, & Davies, 1979).

A study by Bradshaw and Wallace (1971) utilized faces made from Identi-Kit, which consists of transparent overlays each displaying a single feature (of which seven features were used; i.e. nose, mouth, eyes, eyebrows, hair, chin shape, and age lines) to investigate whether faces are processed analytically by their individual features, see Figure 1.1 for examples of stimuli. Participants were presented with 5 pairs of identical faces and 40 pairs of faces differing in two, four, or all seven features presented in both photographically positive and negative forms and had to give a verbal response of whether each set of paired faces were "same" or "diff." It was found that when the number of identical critical features decreased, participants' recognition accuracy decreased and the response time increased. Likewise, inverted faces resulted in more errors and slower processing as compared with upright faces. However, photographically negative faces were processed in a similar manner as photographically positive faces. In the second part of the study, participants were given a practice trial to familiarize themselves with one of each kind of stimulus (i.e. upright and inverted faces in photographically positive and negative forms). Participants were then presented with 40 pairs of identical faces and 5 pairs of faces differing in all seven features and had to give a verbal response to decide whether each pair of faces was "same" or "diff." Consistent with the results found in the first study, participants processed photographically negative and positive faces in a similar manner. Results also showed that reaction time was faster when participants received practice in the second study than without training in the first study, and that having a practice trial greatly decreased the reaction time for the more difficult, inverted faces but not the familiar, upright faces. These results suggest that features may be processed independently as participants' recognition accuracy and reaction time is dependent on the number of crucial features manipulated.

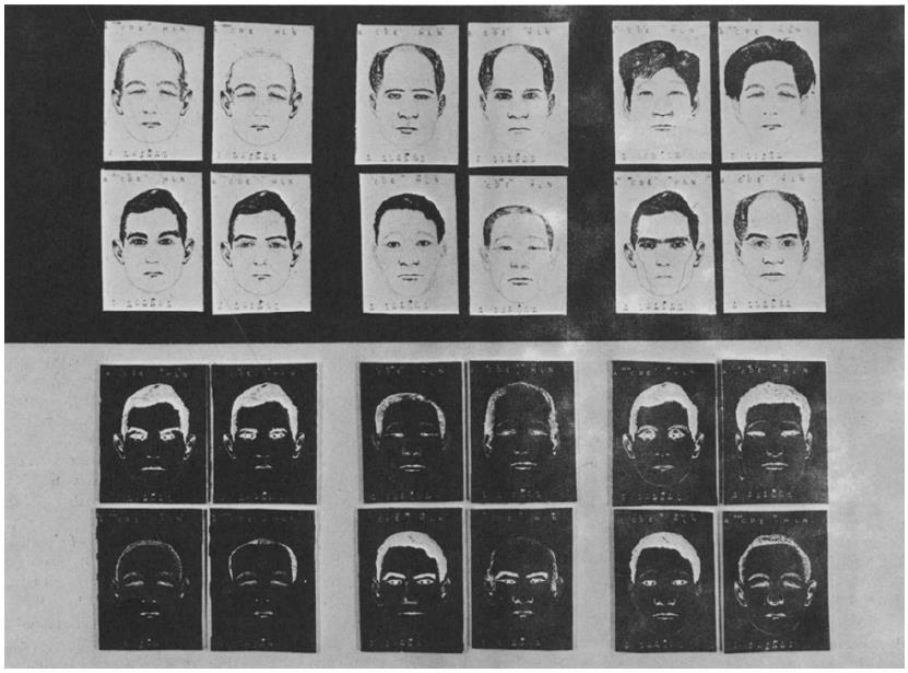


Figure 1.1. Examples of paired Identi-Kit faces differing in two, four, and seven features in both photographically positive and negative forms used in Bradshaw and Wallace's (1971) study.

Contrary to Bradshaw and Wallace's (1971) findings, another study employing a similar same-different judgement task yielded different results (Sergent, 1984). In this study, participants were presented with photo-fit faces that differed in three aspects (i.e. the eyes and eyebrows, the inferior contour, which is the chin and jaw, and the internal space, which is the arrangement of the inner features within the face) in two ways (i.e. round (E1) or large (E2) eyes and eyebrows, V (C1) or U (C2) shaped inferior contour, and large (I1) or small (I2) internal space), see Figure 1.2 for an example of the eight stimuli used. Participants were tested in two blocks: the dissimilarity judgement and the reaction time task. Participants were presented with pairs of different faces and had to rate the dissimilarity between two faces from a scale of 1 to 25 in the dissimilarity judgement task. In the reaction time task, participants were shown the same set of faces in upright and inverted conditions and

had to indicate whether the paired faces were the "same" or "different." Results from the dissimilarity judgment task suggest that participants individuate faces by assessing the relationship between the features to create an overall impression of a face, rather than analytically assessing the individual features of a face. As for the reaction time task, results showed that when the number of differences between stimuli increased, participants' error and reaction time decreased. Further analyses showed that when only one feature was manipulated, the reaction time for recognizing different contours was faster than those of the eyes and internal space for all participants. If faces were processed analytically by independent features, faster detection of a difference in contour should not be affected by additional differences in the eyes or internal space. However, participants' reaction time decreased as the number of differences increased, indicating that faces are not necessarily processed analytically, but rather holistically.

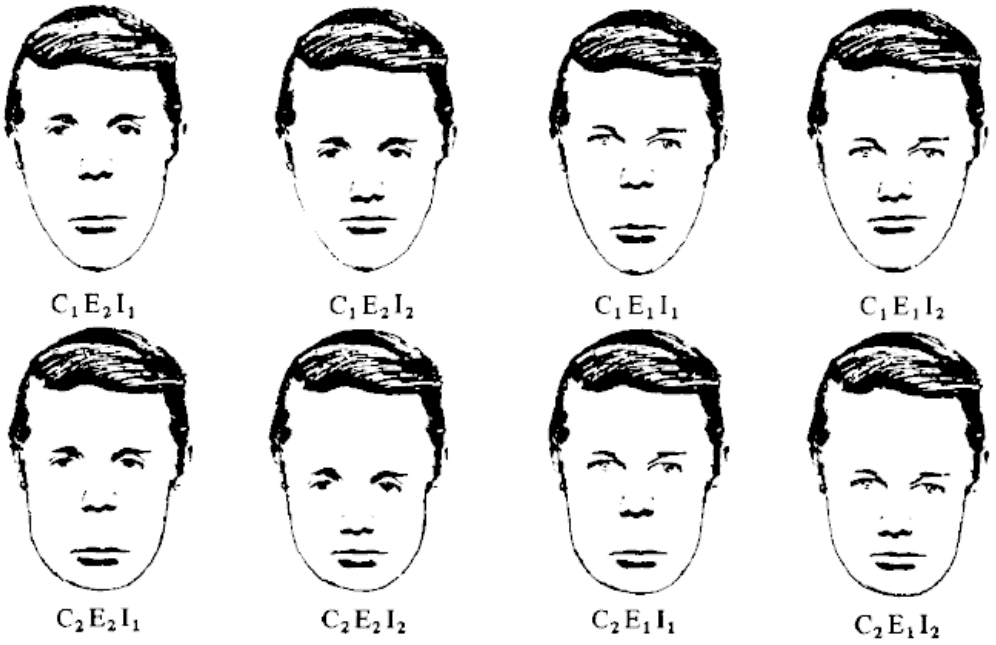


Figure 1.2. Examples of the 8 photo-fit faces used in Sergent's (1984) study; C = contour, E = eyes, I = internal space.

1.1.2. Configural Face Processing Theory

The configural face processing theory proposed that it is more advantageous for observers to perceive a face as a whole configuration by examining the spatial relations and relative positions of the features on the face (e.g. symmetry, the distances between the eyes and mouth or the position of the nose in relation to the face) than to analyze individual features by parts (Rhodes, 1988). The configural processing theory comprised four main concepts: 1) visual objects are processed both holistically/configurally and analytically/featurally, 2) these two processes interact with one another, 3) holistic processing has an advantage over analytical processing, and 4) the efficiency of holistic processing is more important in face perception than other object perception (Bruyer, 2011).

Some studies supported the configural face processing theory using evidence from prosopagnosia (i.e. an inability to recognize faces) and alexia (i.e. an inability to read) cases. Researchers have proposed that faces are processed holistically by examining the overall structure of the face whereas reading requires an analytical processing of the parts by recognizing the letters that make up each word, and words that make up each sentence. Neuroimaging studies have also shown that face and word processing utilize different regions of the brain, providing evidence that the fusiform face area and the visual word form area are specialized for face and word identification respectively (Kleinschmidt & Cohen, 2006). Although holistic processing of faces is disrupted in prosopagnosia, analytical processing of the parts, like reading, would not be affected. On the other hand, in alexia, analytical processing of text is affected but holistic processing required in face recognition remains unimpaired (Kleinschmidt & Cohen, 2006). Numerous face processing studies have

also provided evidence for the configural face processing theory by investigating factors such as the face inversion effect, composite face effect, and part-whole effect.

1.1.2.1. Face Inversion Effect

Farah, Wilson, Drain, and Tanaka (1998) conducted a series of experiments using whole and part masks to examine how participants perceive faces, objects, and words. In their first experiment, participants were presented with pairs of upright or inverted faces, followed by a probe word (i.e. eyes, nose, mouth, or all). Participants were then required to decide whether the probed feature or the whole face of the two faces were the same. Recognition accuracy was greater when faces were presented upright than inverted, indicating that participants processed faces holistically. Additionally, recognition accuracy was greater when irrelevant features were compatible with probed feature than when they were incompatible, providing additional support that faces are processed holistically as the incompatibility of other irrelevant features had a negative impact on participants' recognition performance.

To investigate whether face recognition requires an analytical assessment of facial parts, as does word recognition, a second experiment comparing face and word processing was conducted (Farah et al., 1998). For the word recognition stimuli, each of the 36 four-letter words was paired with an identical copy of the same word and with another word that differed by a single letter (e.g. "most" and "must"). Nine four-letter whole masks (e.g. "wink") and part masks (i.e. scrambled letters of the whole mask, such as "wkni") were also created. As for the face recognition stimuli, each of the 36

faces was paired with an identical copy of the same face and with another similar face that shared the same facial outline (i.e. hair, ears, chin, and shoulders) but differed in all three internal features (i.e. eyes, nose, and mouth). Nine additional faces were used as whole masks and nine part masks were created by scrambling the positions of the features of the face (e.g. nose and mouth placed in the positions of the left and right eyes), see Figure 1.3 for an example of whole and part masks. In each condition, participants were presented with a target stimulus, followed by a whole or part mask. Participants were then presented with another stimulus (either identical to or different from the target stimulus) and had to decide if it was the same or different. Participants were better at recognizing different than same word pairs while participants were better at recognizing identical than different face pairs. There was also a trend towards whole masks impairing recognition for both stimuli more so than part masks, and that this trend was more evident in faces than words. This finding suggests that faces do not require a part based representation during perception as much as words do. However, the authors acknowledged that the differences between face and word stimuli may be due to other methodological factors, such as the size differences between both stimuli or the artificial nature of the part face masks.

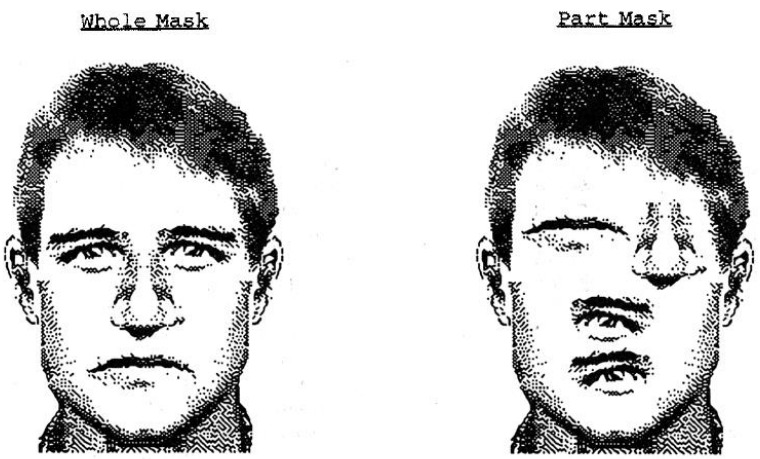


Figure 1.3. Examples of whole and part masks in Farah, Wilson, Drain, & Tanaka's (1998) study.

To control for size differences and the unrealistic nature of part face masks, experiment two was replicated with an additional stimulus type (i.e. inverted faces). Inverted faces and upright faces are identical in size and would have similarly unrealistic part masks. Therefore, if the effects of masks on face and word recognition were due to low-level methodological differences, masking upright and inverted faces should produce similar results. On the other hand, experiment one has shown an advantage for recognizing upright over inverted faces; if the effects of face masks were due to face-specific perceptual representation, performance may differ for upright and inverted faces. Results revealed that perceiving upright faces was more affected by whole than part masks; however, words and inverted faces did not show the same pattern. This finding provided evidence that the differential effects of masks were not due to low-level methodological differences; whole face masks may have impaired recognition accuracy for upright faces because of face-specific perceptual representation (Farah et al., 1998).

To determine whether the results of previous studies extend to other concrete objects, stimuli of houses were presented in addition to words and upright faces. Results showed that face perception was most affected by whole masks, followed by house perception, then word perception. This suggests that perception involves a continuum of representation, with faces being represented holistically, words analytically, and objects falling in between the continuum (Farah et al., 1998).

1.1.2.2. Composite Face Effect

In a study by Young, Hellawell, and Hay (1987), composite faces were created (by combining top and bottom halves of different faces) to examine how observers process and recognize faces. The top and bottom face halves of two well-known politicians were combined and participants were required to name the face based on the cued half of the face while disregarding information from the other half. Participants had difficulty directing attention to the cued half only, suggesting that recognition is not based on the individual features of the face but rather a construct of face representation.

A recent eye tracking study with new-borns, 3-month-old infants, and adults also found that participants employed a holistic processing style when passively viewing photos of aligned or misaligned composite faces, see Figure 1.4 (Turati, di Giorgio, Bardi, & Simion, 2010). All participants were presented with only the top half of a face in the habituation-familiarization phase. During the following two test phases, pairs of composite faces were presented in either the aligned or misaligned condition. Within each pair, the bottom halves of the faces were the same; one of which the top half of the face

was identical to the top half face shown in the habituation-familiarization phase, while for the other the top half was new. Eye tracking results showed new-borns spent similar amount of time viewing familiar and novel faces in both conditions, implying that new-borns were unable to discriminate between familiar and novel faces. Three-month-old infants focused more on the top half of the face for both conditions. When faces were aligned, 3-month-olds paid equal attention to both familiar and novel faces, implying that newly-combined faces had a negative impact on infants' ability to discriminate between the two composite faces; conversely, 3-month-olds fixated longer at familiar top half faces when faces were misaligned, suggesting that infants were able to discriminate between the two stimuli in the misaligned condition. Finally, eye tracking results with adults revealed a similar pattern as 3-month-old infants; adults showed a preference for the top half of the face for both conditions and in the aligned condition, adults directed equal attention to both familiar and novel faces, indicating that the bottom half of the face also impaired adult's ability to discriminate between familiar and novel faces. In the misaligned condition, however, adults tended to pay more attention to novel rather than familiar top half faces, possibly because it was unnecessary for adults to re-examine familiar faces in the test phase. The findings of this study suggest that holistic face processing occurs as early as 3-month-old.

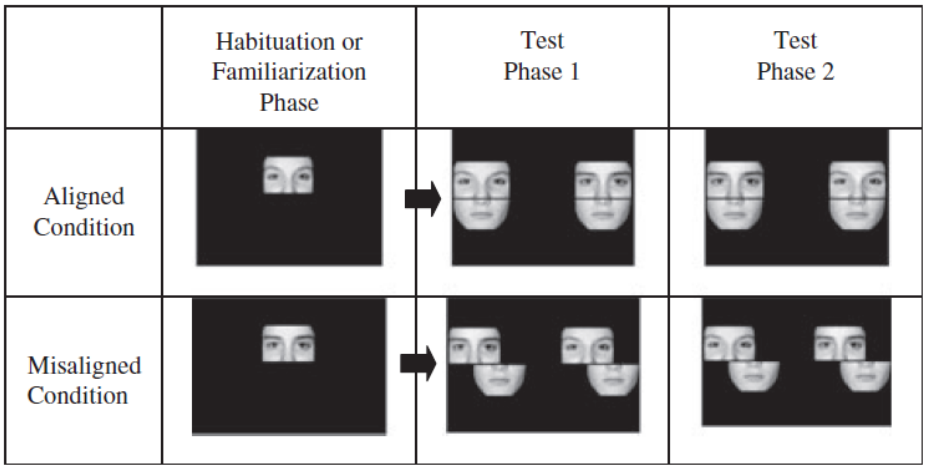


Figure 1.4. Examples of (aligned and misaligned) stimuli and procedure used in Turati, di Giorgio, Bardi, &Simion's (2010) study.

1.1.2.3. Part-Whole Effect

Tanaka and Farah (1993) conducted a series of experiments to investigate whether observers recognize faces analytically by their isolated parts (i.e. individual features) or holistically. In the first experiment, participants were presented with scrambled and intact faces along with their names spoken verbally. During the test phase, participants were required to identify one of the features (i.e. eyes, nose, or mouth) in one of three conditions (i.e. isolated part of the feature only, scrambled face, or intact face), see Figure 1.5 for an example of the stimuli. Participants had to identify each feature twice; once in isolation and once in the context of a whole face. Results showed that participants were better at identifying features from intact faces in whole than in isolation; additionally, participants were better at identifying features from intact than scrambled faces when presented in whole faces and isolated parts. This suggests that intact faces are recognized using a more holistic representation than scrambled faces (Tanaka & Farah, 1993).

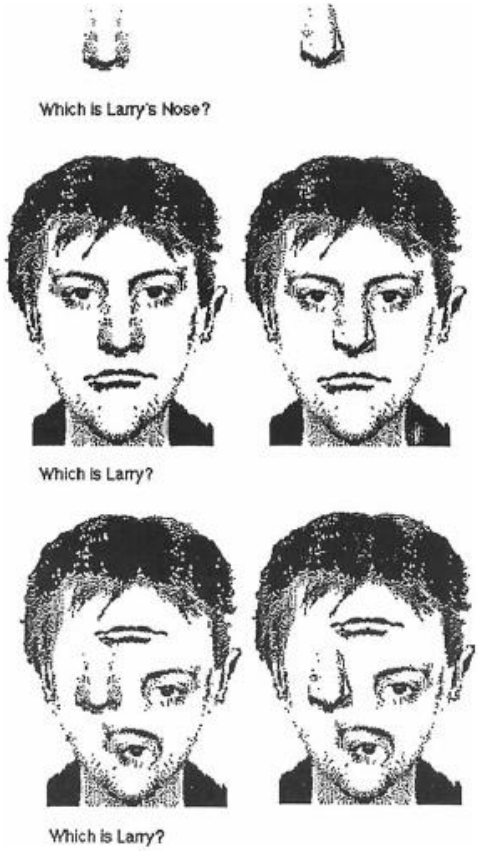


Figure 1.5. Examples of isolated parts, intact faces, and scrambled faces in Tanaka & Farah's (1993) study.

Because scrambled faces may be considered unnatural as compared with normal faces, another experiment with inverted faces was conducted (Tanaka & Farah, 1993). Similar with the procedure of the first experiment, participants were presented with upright and inverted faces along accompanied by their names spoken verbally in the learning phase. Participants were then required to identify the features in whole faces and isolated parts. Results showed that participants were better at recognizing the features of upright faces when presented in whole faces than in isolation; however, there was no advantage for whole faces over isolated parts for inverted faces. Although identical visual information was presented in both

test conditions for both face orientations, there was an advantage for whole faces over isolated parts when faces are presented upright, but not inverted. The disruption to face recognition performance for inverted faces provided support to the hypothesis that faces are represented holistically.

To examine whether holistic processing is special to face perception or can be extended to objects, the third experiment compared face and object (i.e. house) recognition. Participants learned images of faces and houses associated with a person's name. During the test phase, participants had to recognize the features of the face (i.e. eyes, nose, or mouth) and house (i.e. door, small window, or big window) in two conditions (i.e. whole-object and isolated part). Face recognition results replicated those found in experiment two; with better recognition of whole faces than isolated parts. As for object recognition, there were no differences between whole objects and isolated parts for houses. This indicates that face perception relies on holistic representations more so than object recognition (Tanaka & Farah, 1993).

1.1.3. Information Processing Models

To understand the human face processing system, Bruce and Young (1986) set out to identify the different types of information, or "codes," derived from face processing and the functional components responsible for generating and accessing these information codes. It was suggested that seven types of information codes are involved in face processing: 1) pictorial, 2) structural, 3) visually derived semantic, 4) identity-specific semantic, 5) name, 6) expression, and 7) facial speech codes, of which the last two do not play a direct role in face recognition. Utilizing these information codes, observers structurally encode faces, producing a set of view-

centered descriptions (used to analyze facial expressions and facial speech) or expression-independent descriptions (used to provide information about the identity of the face).

Bruce and Young's (1986) identified three stages involved in face recognition that progresses sequentially, see Figure 1.6 for a visual representation of the model. After structurally encoding the visual information about a target face, the first stage involves the matching process, in which the target face is compared with a set of visual descriptions previously stored, known as *face recognition units* (FRU). FRUs allow a familiar face to be distinguished from other faces and the unique FRU associated with a particular person becomes active when the appropriate face is shown. The activation of an FRU leads to the second stage, in which the memory of semantic information about a person is activated. These units, known as *person identity nodes* (PIN), allow an observer to recognize someone as a "person" rather than by their face and, thus, can be activated by inputs such as one's voice or (written or spoken) name other than one's face. After activation of an appropriate PIN, the last stage would be name generation.

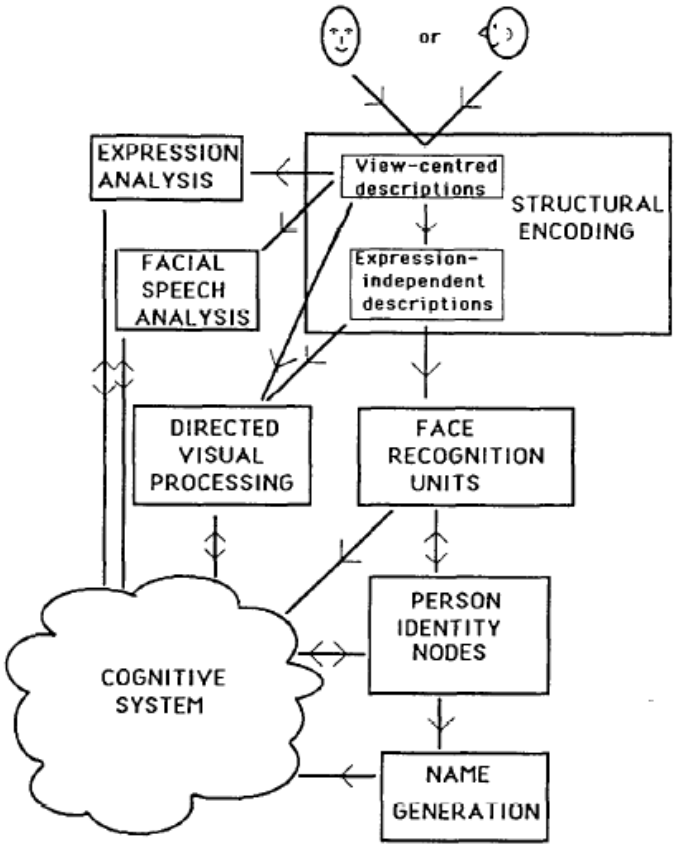


Figure 1.6. A functional model of face processing proposed by Bruce and Young (1986).

Based on the basic concepts of Bruce and Young's (1986) face recognition model, Burton, Bruce and Johnson (1990) proposed an interactive activation model that clustered information codes into several pools of units: *face recognition units* (FRU), *personal identity nodes* (PIN), *semantic information units* (SIU), and *name input units* (NIU). All units are connected within a pool and across pools, see Figure 1.7 for a visual representation of the model. When one of the units receives input, information is passed around the system and may activate another pool of units. The interactive activation model propose a bi-directional link between the pools of units and suggests that the person, rather than the face, is recognized and decisions about familiarity are made at the person identity nodes (PIN).

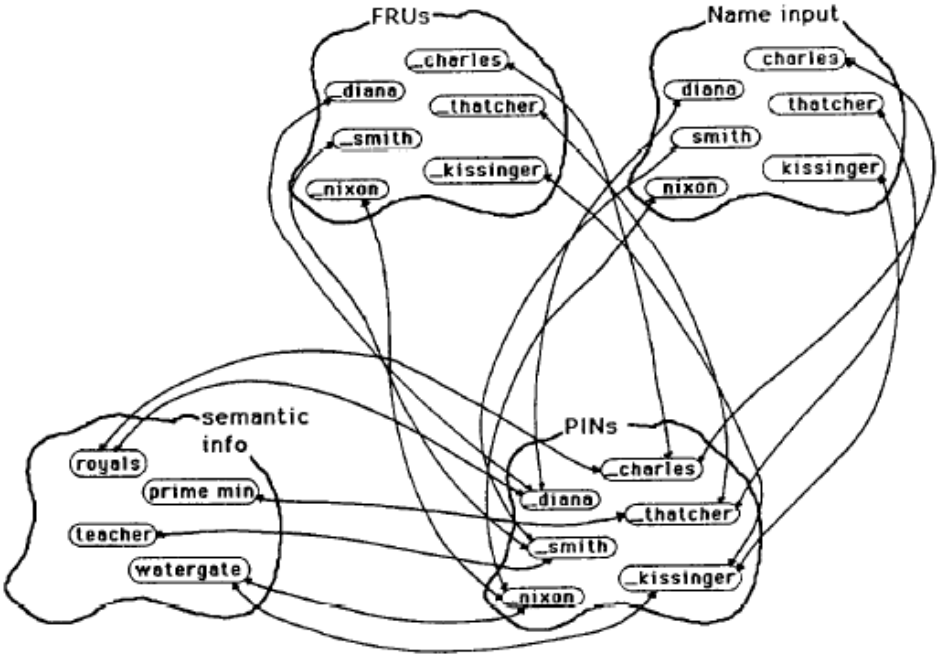


Figure 1.7. The interactive activation model of person recognition proposed by Burton, Bruce, and Johnson (1991).

Case studies involving prosopagnosic patients have provided evidence supporting selective components of the face recognition models. Prosopagnosia is an inability to recognize faces due to damage to one's temporal lobes (Wickens, 2009). Individuals with prosopagnosia are unable to recognize familiar faces despite being able to identify people by voice or to recognize people's age, gender, and even facial expressions. For example, a study with a prosopagnosic patient (P.H.) revealed that an intact semantic memory allowed accurate recognition of familiar names but damage to face recognition units resulted in an inability to recognize familiar faces overtly (De Haan, Young, & Newcombe, 1987). P.H. was presented with face-name interference stimuli in three task conditions: a face was either paired with the correct name, a related name, or an unrelated name, see Figure 1.8. P.H. was required to categorize the names (i.e. politicians or nonpoliticians) while ignoring the distractor faces. Reaction time for the unrelated condition was longer than reaction times for the

related and same person conditions, implying that although P.H. could not overtly recognize faces, the interference of distractor faces indicate that P.H. could covertly recognize and access semantic information of familiar politician faces.

Damage to the face recognition units (i.e. first stage of Bruce and Young's face recognition model) did not interfere with one's ability to retrieve semantic information (i.e. second stage) associated with a name (i.e. third stage), supporting Bruce and Young (1986)'s account that PINs can be accessed by other input, such as one's written name. Although damage to the FRU had a negative impact on overt recognition abilities, P.H. showed covert recognition with the interference effect, suggesting that prosopagnosic patients are still able to use representations of faces that may not be accessible overtly. Young and De Haan (1988) proposed that P.H.'s inability to retrieve semantic information associated with a face may be due to a disconnection of FRU outputs, leading to damage to the pathways between FRU and PIN. However, it does not explain how a disconnected FRU can continue to influence name recognition.

Conversely, Burton et al. (1990)'s interactive activation model proposed that the interference effect shown by P.H.'s semantic priming from distractor faces is affected by the connection strengths between the pools of FRUs and PINs. When an observer perceives a face, an active FRU passes on information to the pool of PIN, which can also be passed on to the SIU; in P.H.'s case, because the connections between the pools of FRUs and PINs are weak, the level of activation in PIN would be smaller than an intact system, but can be passed on to the pool of SIU normally. The below threshold activation of the appropriate PIN may not allow for overt recognition of faces but may cause interference as reflected by P.H.'s semantic priming from unrecognized faces. These findings suggest that examining overt

responses alone may not be adequate to determine what is preserved or impaired in the cognitive processing of faces by prosopagnosic patients (Bruyer, 1999). Moreover, covert recognition has only been identified in some, but not all, prosopagnosic patients, which suggests that prosopagnosia may have different forms that should be distinguished (Bruyer, 1999; Newcombe, Young, & De Haan, 1989).

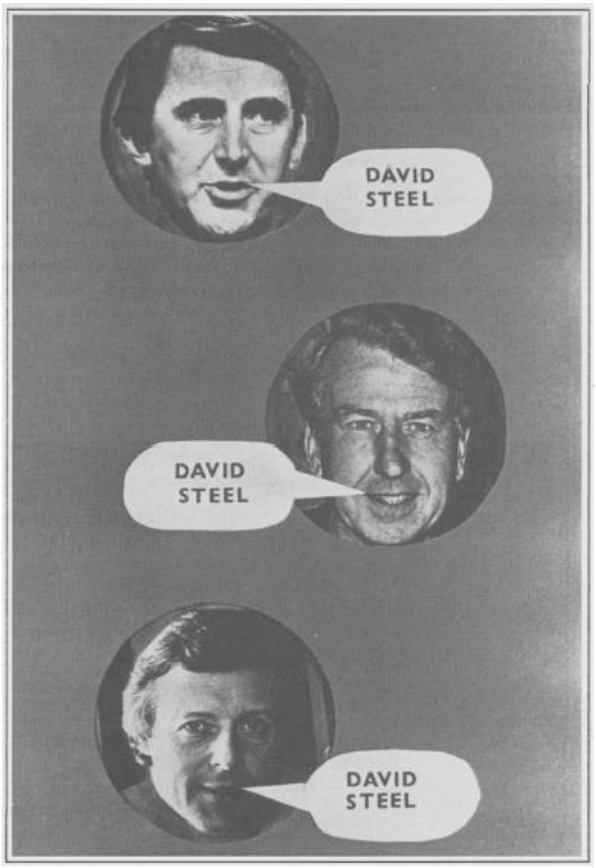


Figure 1.8. Examples of face and name stimuli from the same person (top), related (middle), and unrelated (bottom) task conditions in De Haan, Young, &Newcombe's (1987) study.

A similar methodology was used with a different prosopagnosic patient, M.S. (Newcombe et al., 1989). M.S. was presented with a face or name prime before having to classify target names of people, who are either famous or not. Each prime was either related to, neutral, or unrelated to the target name. M.S. responded quicker

when presented with related name primes than neutral ones, but there were no differences between the reaction time for neutral and unrelated name primes. M.S. showed no priming from faces he could not recognize overtly, indicating that covert recognition is not evident across all cases of prosopagnosia (De Haan et al., 1987). The contrasting findings with P.H. and M.S. may be a result of the severity of M.S.'s high-order perceptual impairment. According to Bruce and Young's model (1986), M.S's inability to construct visual representations of faces and objects affected the structural encoding process, preventing M.S. from achieving overt and covert recognition of faces. On the other hand, the interactive activation model (Burton et al., 1990) proposes that the link between FRUs and PINs may have been damaged in M.S.'s case, preventing activation at the PIN level when a face was presented (Burton, Young, Bruce, Johnston, & Ellis, 1991). Moreover, input to and output from the FRUs must have been impaired, which prevented input that activates the FRUs and output that passes on information to another pool of units (i.e. PINs).

1.2. The Theoretical Framework of Facial Expressions of Emotion

The mutual understanding of emotion through the exchange of facial expressions is one of the most vital aspects of human social interaction. Although facial expressions of emotion are long thought to be universal (Darwin, 1872; Ekman & Friesen, 1971; Matsumoto & Willingham, 2009), cross-cultural studies have found conflicting evidence that challenges the universality hypothesis (Birdwhistell, 1970; Klineberg, 1940; Biehl, Matsumoto, Ekman, Hearn, Heider, Kudoh, & Ton, 1997; Moriguchi, Ohnishi, Kawachi, Mori, Hirakata, Yamada, Matsuda, & Komaki, 2005).

1.2.1. The Universality Hypothesis

In the 1870s, Charles Darwin hypothesized that facial expressions of emotions are universal. His early research involved Englishmen living or traveling in Africa, America, Australia, Borneo, China, India, Malaysia, and New Zealand. Participants reported that the facial expressions of emotion they saw in these countries were the same as their own, leading Darwin to conclude that all humans display the six basic emotional states ("happy," "surprise," "fear," "disgust," "anger," and "sad") using the same facial movements that are biologically and evolutionarily determined. He argued that the meanings of facial expressions would have diverged if they were merely cultural traits that are socially learned (Darwin, 1872).

Numerous studies have provided evidence in support of the universality hypothesis thereafter. Ekman and Friesen (1971), for instance, studied the categorization of facial expressions in New Guinea, a preliterate culture that has minimal exposure to the same mass media presentation of facial expressions in literate cultures. Participants were told a story, followed by the presentation of a set of three faces, and were asked to choose a face showing the emotion that corresponds to the story. Results showed that New Guinean adults and children could accurately categorize most of the facial expressions, but have difficulty distinguishing between "fear" and "surprise" faces. Ekman and Friesan (1971) concluded that facial expressions are universally linked with particular emotions and accounted for New Guineans' confusion of "fearful" and "surprised" faces by stating that the two facial expressions were poorly judged because of the cultural interpretation of fearful situations as surprising. Hence, New Guineans miscategorize "fear" as "surprise," not because of their inability to distinguish between the two emotions but because in their culture, fearful situations are almost always regarded as surprising (e.g. to be taken by

surprise by an unexpected appearance of a ghost, sorcerer, or an aggressive person from another village would be regarded as fearful situations).

A recent study by Matsumoto and Willingham (2009) also supported the universality hypothesis by examining the spontaneous display of emotions of blind individuals. The expressions displayed by congenitally blind and noncongenitally blind athletes in the 2004 Paralympic Games and sighted athletes in the 2004 Olympic Games were compared by examining the facial emotion configurations and individual facial actions produced. Results showed no difference in the facial expressions of emotion between the three groups, providing evidence in support of the hypothesis that the display of emotions is not learned and modelled from observation but are genetically coded for all humans.

1.2.2. The Culture Specificity Hypothesis

On the other hand, researchers who hold a culture specific view believe that facial expressions of emotion vary according to culture. Birdwhistell (1970), for example, proposed that the display of emotion is composed of units and organization similar to the components of a spoken language. Hence, facial expressions and body movements would be learned according to the structure of individual society. Klineberg (1940) held a similar view that facial expressions of emotion, like language, acts as a means of communication and must be socially learned. Although certain patterns of behavior, such as laughing and crying, are universal, the expressions of emotion like “anger” and “fear” are not. In Klineberg’s research, he hypothesized that expressions of humans and another primate should closely resemble each other if facial expression was an innate, evolved behavior. However, he found that humans

were inaccurate in judging a chimpanzee's facial expressions, leading him to conclude that human expressions must be a result of socially learned behavior.

In a study by Biehl, Matsumoto, Ekman, Hearn, Heider, Kudoh, and Ton (1997), a total of 501 participants from the United States, Poland, Hungary, Japan, Vietnam, and Sumatra viewed 56 Japanese and Caucasian facial expressions of emotion and classified them into seven categories ("anger," "contempt," "disgust," "fear," "happiness," "sadness," and "surprise"). Participants also rated the intensity of each expression on a 9-point scale during their second viewing. Results revealed that while observers from all cultures agreed on the facial expressions of emotion for each photograph, the level of agreement between each culture was considerably different. For example, the Japanese had less agreement than other countries on the expressions of "anger," "fear" and "sadness," whereas Americans agreed less than other countries on "contempt." This finding implies that the semantic or affective meanings of each emotion may vary across culture. Further examination of observers' intensity ratings for the 56 images also showed cross-cultural differences. Western observers (i.e. participants from Hungary, Poland, and the U.S.) gave higher intensity ratings for all eight happy expressions while non-Western observers (i.e. participants from Japan, Sumatra, Vietnam) rated seven of eight fear expressions higher than Western observers did. Interpretation of intensity ratings for other emotions was not as direct; with Westerners giving higher intensity ratings for four out of eight expressions of anger, contempt, and surprise, and three of disgust and sadness, whereas non-Western observers gave higher intensity ratings for two of eight expressions of anger, and one of contempt. These findings suggest that there may be dissimilarities in either the portrayal or judgment of facial expressions, or perhaps intensity ratings may be

positively correlated with participants' confidence level in the categorization of expressions.

Recent research continues to provide supporting evidence for the cultural variation view. In a brain imaging study by Moriguchi et al. (2005), 32 Caucasian and Japanese participants observed 16 photographs of "happy" and "fearful" own- and other-race faces. The activation of different regions of the brain was measured by fMRI and results showed that Caucasian and Japanese participants responded differently to "fearful" faces. Caucasians showed a more direct and emotional response with greater activations of the left posterior cingulate cortex (which has been shown to be activated when participants are experiencing threat-related stimuli), the right supplementary motor area (which could affect one's limbic and motor systems when experiencing negatively emotional stimuli), and the paracentral lobule (which has been shown to respond to increased visual attention, suggesting that Caucasians may be directing their attention to fearful faces more than Japanese participants). Japanese, on the other hand, identified facial expressions more systematically by activating the right inferior frontal gyrus (which was suggested to act as a template matching system that matches the target facial expression to previous prototypes of facial emotions in the brain) and the right dorsal premotor area. The differential neural basis in emotional recognition suggests that the perception of facial expressions of emotion may be influenced by individual culture.

1.2.3. Ekman's Theoretical Framework

Ekman (1970) attempted to reconcile the two extreme ends of the argument by differentiating elements of facial expressions that are universal and culture specific. He proposed that the expressions of emotion might vary across cultures as a result of

three main factors: 1) in different cultures, certain stimuli are learned to be the elicitors of specific emotions, 2) different cultures may have different display rules for specific social situations, and 3) the behavioral consequences of facial expressions of emotion may vary across cultures.

Although some evoking stimuli are universally linked with a specific emotion, various stimuli elicit different emotions based on each culture's social practice. Therefore, the comparison of a same event in two cultures may not be accurate, as it does not necessarily elicit an identical emotion. A husband's second marriage, as an example, may evoke a wife's jealousy in one culture but may be a stimulus for pride in another. Rather than assuming that the facial expressions of a wife witnessing her husband's second marriage vary across cultures, it should be recognized that the stimulus of a "second marriage" evokes different emotions in different cultures.

Display rules, which are socially learned techniques, may also govern one's display of emotions in particular social settings. Ekman (1970) identified four management techniques that may be used to control one's facial expressions: 1) to de-intensify one's emotions, 2) to over-intensify one's emotion, 3) to look neutral, and 4) to mask one's true emotion with another emotion. These display rules are culture specific and determine the appropriateness of one's facial expressions in varying social situations. For example, even if the same event evokes joyous emotion in two cultures, it is likely that the display rules for the emotion will be different; one culture may de-intensify joy with a modest smile while another may over-intensify the emotion by jumping and shouting in joy.

Besides the cultural differences in evoking stimuli and display rules, the behavioral consequence of an emotion is another factor that may differ across cultures.

The body postures and movements that occur as a result of experiencing a certain emotion is one observable behavioral consequence of an emotion. Although the body postures and movements can be interpreted as a coping mechanism associated with the emotion, these behavioral consequences do not distinguish between emotions. As an example, an individual may choose to walk away from a situation to cope with overwhelming feelings of anger, fear, or even disgust. Thus, observers may not be able to correctly identify the emotion felt by an individual based on one's body movements alone as the behavioral consequence does not necessarily reflect the actual emotion felt.

Ekman (1970) argued that although the evoking stimuli, display rules, and behavioral consequence of emotion may differ across cultures, the facial expressions of the primary emotions are universal to mankind.

1.3. Ecological Validity

Face processing is an important aspect of our lives. Recognizing close friends and family or unfamiliar acquaintances and neighbors in our daily lives, helping police identify culprits in a crime, or noticing the emotional state of a person during a conversation are a few examples of real world situations that require our ability to recognize identities and emotions. Although face recognition studies showed high performance level in laboratory settings, accurate identification of faces does not seem to be replicable in some real life situations, such as eyewitness testimony.

1.3.1. Eyewitness Testimony

Even though eyewitness testimony has been the leading source of evidence that jurors find most convincing when deciding the verdict in criminal trials, research

has found that eyewitness misidentification is the primary cause of wrongful convictions, with an estimate of 4,500 cases of mistaken convictions each year in the United States (Pezdek, 2006; Wells & Olson, 2003). Various factors, including estimator variables (e.g. characteristics of the witness, event, and testimony) and system variables (e.g. pre-line-up instructions, line-up content, and line-up presentation method), have been found to affect the recognition accuracy of eyewitness identifications. Research revealed that the transformation of stimuli, a change of perspective and presentation method, which would naturally occur from laboratory studies to real life situations, may reduce recognition accuracy. Therefore, it is crucial for researchers to develop and utilize naturalistic stimuli that accurately represent a typical social situation to produce results which are more ecologically valid.

Eyewitness testimony research has aimed to identify whether certain groups of people are better eyewitnesses than others. For instance, studies examining the effect of race found a robust phenomenon (known as the other-race effect) across various experimental conditions, with better recognition accuracy for own- than other-race faces (Meissner & Brigham, 2001; see a more detailed discussion of the other-race effect in Chapter Three). The other-race effect is suggested to be a reflection of an individual's racial experience; a bias which may be reduced with increased exposure or familiarity with the other races (Valentine, 1991; Hancock & Rhodes, 2008). A range of research has been conducted to understand how observers perceive faces, to discover ways of reducing the other-race effect, and to improve observers' recognition abilities because of its severe implications in eyewitness testimony.

A study examining the relationship between contact level, configural coding, and recognition performance has found that higher levels of self-reported contact with

other-race faces was associated with a reduced other-race effect in recognition performance and configural coding for both Chinese and Caucasian participants (Hancock & Rhodes, 2008). Although higher quantity and quality of contact were found to be associated with decreased other-race effect, the quantity of contact is not necessarily negatively correlated with the other-race effect as studies have found that living in a culturally diverse environment does not sufficiently reduce the other-race effect. Instead, quality interaction, which requires one's ability to differentiate other-race faces, provides the opportunity for observers to learn the diagnostic features that individuate faces of other-races (Hills & Lewis, 2006).

In another study, Johnson and Fredrickson (2005) found that positive emotion can reduce the other-race effect. Watching videos that elicited feelings of joy, as compared with fear and neutral feelings, prior to face learning or testing improved Caucasian participants' ability to recognize African faces, suggesting that positive emotions may reduce intergroup distinctions and promote a more inclusive social categorization, which in turn facilitates more accurate memories of faces from the other races. However, negative emotion, such as anger, was also found to reduce the other-race effect (Ackerman, Shapiro, Neuberg, Kenrick, Vaughn Becker, Griskevicius, Maner, & Schaller, 2006). When presented with neutral Caucasian and African male faces, Caucasian participants showed the other-race effect; however, recognition accuracy was higher for angry African than Caucasian faces. This suggests that threat cues displayed by other-race faces may elicit observers' fear of danger, which in turn facilitates self-protection or other mechanisms that allow differentiation between people of other-races as they may pose greater threat than people of their own-race.

1.3.2. Methodological Considerations

The incongruence between recognition performance in laboratory studies and real life settings may be due to the artificial nature of laboratory studies. Most face processing studies in laboratory settings involve static stimuli, such as illustrated emoticons (e.g. Yuki, Maddux, & Masuda, 2007) or photographs of posed facial expressions (e.g. Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Jack, Blais, Scheepers, Schyns, & Caldara, 2009) or natural faces (e.g. Matsumoto & Willingham, 2009), as they allow greater control over the development and presentation of stimuli (e.g. standardization of how the images were photographed and processed or the time and distance the stimuli was shown to participants). However, utilizing static stimuli may not be an accurate representation of real life situations as evident in eyewitness studies mentioned earlier.

Firstly, recognizing two-dimensional images of faces may not be the same as recognizing real faces that offer more visual information from a three-dimensional view. In a typical laboratory study, for instance, participants are asked to recognize a person by the same photograph shown in the learning phase for a considerable amount of time; however, in a real life setting such as eyewitness identification, the eyewitness may only have a glimpse of the culprit from a particular point of view but would be required to identify the culprit in a different manner (e.g. line-up or photos from a different viewpoint), which considerably increases the difficulty of the task. Besides eyewitness studies that found a negative impact of transformations of stimuli (e.g. intentional disguises or natural changes in appearance) on eyewitness identification performance, face processing studies have also found that rotating faces in depth during the learning and recognition phase may impair participants' recognition accuracy (Hockley et al., 1999; Read et al., 1990; Valentin, Abdi, &

Edeleman, 1999; Hill, Schyns, & Akamatsu, 1997; O'Toole, Jiang, Roark, & Abdi, 2006).

Presenting a face from different rotations in depth (e.g. 45-degree or 90-degree) produce different views of the face; for an observer to recognize a face from a different viewpoint than the initial view that the face was learnt, an observer would need to encode information that is unique to the face in order to distinguish the target face from all other faces and to retrieve that unique information to identify the target face from a new point of view (Hill et al., 1997; O'Toole et al., 2006). How humans achieve a representation of faces for successful recognition from various viewing conditions is an intriguing question that researchers seek to answer. Two opposing theories speculate how faces may be represented in human minds. Part-based (or structure-based) theories propose that a three-dimensional view of a face is structured by analyzing individual parts of the face; hence, observers create a representational view of a face by its parts and components, independent of the original image of the face that was presented from a specific viewpoint. On the other hand, view-based (or image-based) theories posit that the representation of a face is constructed from a direct analysis of multiple images, which is then stored in memory as a collection of multiple views; therefore, familiarizing oneself with different views of a face will enhance an observer's ability to recognize the face from a new viewpoint (Hill et al., 1997; O'Toole et al., 2006).

Studies have found that rotating faces in depth impairs recognition accuracy (O'Toole et al., 2006; Valentin et al., 1999). Changing the orientation of a face between 0-degree and 90-degree during the learning and recognition phase decreases participants' recognition accuracy and increases participants' response bias and response time, suggesting that it is more challenging for observers to recognize faces

after a change of pose (Valentin et al., 1999). Studies have also discovered an improvement in recognition performance when a face is presented from a three-quarter view (i.e. a 45-degree rotation in depth), possibly because a three-quarter view presents the most diagnostic information that is crucial in face identification (Hill et al., 1997; O'Toole, et al., 2006). These findings provide evidence that a change of perspective and presentation method (a change that would occur from laboratory studies to real life situations) may, in fact, reduce recognition accuracy.

Additionally, recognizing static images may be different from recognizing dynamic identities and emotions because real faces are dynamic. Taking expressions of emotion as an example, real emotion may unveil subtly and gradually during a social interaction, revealing ample information about one's external and internal state through facial motions or head movements; however, posed photographs do not provide any dynamic information at all as they only capture a single image of an intense emotion. Therefore, utilizing dynamic stimuli in face processing studies would provide a more accurate representation and natural depiction of a typical social encounter and may unveil results different from studies utilizing static stimuli.

Facial movements (which can be identified as either rigid or non-rigid motions) provide different views of a face for an observer to better construct a more complete representation of the face and derive social information about the person. Rigid motions, such as nodding, shaking, tilting, and rotating of the head, provide social interaction signals to another person. For example, turning of head towards or away from someone can signify the beginning or ending of a conversation, nodding of the head shows agreement, and shaking of the head may indicate disapproval or disagreement. Non-rigid motions, on the other hand, can be categorized into facial expressions, eye gaze changes, and facial speech. Some non-rigid motions, like facial

expressions, may distort the identifying features of a face. For instance, the distances between features would be completely different when a person is surprised (e.g. raised eyebrows, wide eyes with mouth ajar) and angry (e.g. furrowed eyebrows, flaring nostrils, and clenched teeth). A previous study exploring the role of head and facial movements on recognition ability discovered that rigid and non-rigid head movements provided useful cues for observers to categorize identity and gender respectively (Hill & Johnston, 2001). Participants may have found rigid head movements more useful for categorizing identities because non-rigid motions were uninformative in providing differentiating cues for identities as all "actors" were performing similar movements (i.e. telling jokes).

Although facial movements convey rich social cues that may provide ample visual information about a face, they may prove challenging to the perceivers as these visual cues not only need to be constantly examined and interpreted, they may also alter the facial structure that is crucial to recognize unfamiliar faces (O'Toole, Roark, & Abdi, 2002; O'Toole et al., 2006). Therefore, researchers question how dynamic stimuli would affect observers' recognition accuracy and how facial motion may be advantageous or disadvantageous in face recognition.

There are two hypotheses about the role of facial motion in face recognition. The representation enhancement hypothesis suggests that facial motion provides observers additional structures of a face, which would enhance the encoding of the invariant structure of a face. Therefore, viewing dynamic stimuli may be advantageous for face recognition as movements allow observers to build a three-dimensional representation of a face (O'Toole et al., 2002; O'Toole et al., 2006; Knappmeyer, Thornton, & Bulthoff, 2003). On the other hand, the supplemental information hypothesis posits that facial motions reflect characteristics that are

specific to an individual, providing additional information to the invariant structure of the face. Studies have shown that an individual's dynamic facial signatures (e.g. one's characteristic crooked smile or a head tilt) may enhance recognition for familiar or famous faces, especially in non-optimal viewing conditions (e.g. blurred, negative or inverted faces); however, this advantage may not be evident in unfamiliar or unknown faces as observers may need time to learn the dynamic facial signatures that would provide reliable identification information (Knight & Johnston, 1997; Lander, Christie, & Bruce, 1999; Lander & Bruce, 2000). Additionally, by presenting an identical amount of static information for both dynamic and static conditions, findings provided evidence that the dynamic movement of a face, rather than the amount of static information, enhances observers' recognition accuracy (Lander et al., 1999). The timing and direction of dynamic stimuli may also affect participants' recognition of famous faces; more specifically, participants were best at recognizing an original forward-moving sequence of images presented at normal speed, whereas reversed presentation of the moving clip and speeding up of the original clip decreased participants' recognition ability (Lander & Bruce, 2000). These findings indicate that facial motions provide additional information that may reflect characteristics specific to an individual, resulting in better recognition of dynamic faces.

1.4. Summary

Although faces are complex in nature, observers are proficient in recognizing the identities or expressions of faces. A wide range of research has been conducted to investigate how faces are perceived and processed to better understand how observers achieve recognition of familiar and unfamiliar faces. Neuroimaging studies have suggested that there is a specialized mental module, possibly located in the fusiform

face area, dedicated to face processing (Kleinschmidt & Cohen, 2006). Case studies with prosopagnosic patients provide evidence to neuroimaging studies as patients with impaired face recognition abilities continue to be able to access memory and process other visual objects without difficulty (De Haan et al., 1987; Burton et al., 1991). Behavioral studies exploring how faces are analyzed suggest that individual features and the configuration of a face are both crucial in face processing. Rather than analytically examining individual features of a face or holistically processing the overall arrangement of features, observers use both types of information interactively. Studies showing the face inversion effect, composite face effect, and the part-whole effect suggest that holistic processing have an advantage over analytical processing. Besides vast research on face processing, different information processing models have also been proposed to account for varying findings about face recognition. Based on Bruce and Young's face recognition model (1986), which suggested seven main information codes that can be processed independently in three progressive stages, Burton, Bruce and Johnson (1990) proposed an alternative interactive activation model that suggested a bi-directional link between pools of units in which information codes are clustered.

Even though previous face processing research has expanded our understanding of how recognition may be achieved, other areas (such as the relationship between culture, eye movement, and face processing tasks such as face recognition and emotion recognition) remain under-explored. Recent cross-cultural studies employing eye tracking methodologies have found systematic differences in the way observers perceive and recognize identities and emotions (Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Jack, Blais, Scheepers, Schyns, & Caldara, 2009). Moreover, study exploring the effects of cultural environment on populations such as

British born Chinese have revealed that exposure to and familiarity with another culture may foster a hybrid fixation pattern that is beneficial for recognizing both own- and familiar other-race faces (Kelly, Jack, Miellet, De Luca, Foreman, & Caldara, 2011a; see Chapter Three for an examination of the data). Therefore, face processing studies utilizing eye tracking that involve participants from unique cultures, such as Malaysia, may provide insight as to how culture may affect an observer's eye movement strategies and recognition abilities. The relationship between culture and face processing tasks, such as face recognition and emotion recognition, will be discussed in Chapter Two.

Chapter Two

Cultural Differences

2.1. The Holistic and Analytic Dichotomy

It has long been thought that cognitive and perceptual processes are similar across cultures. However, evidence from cross-cultural studies, especially those comparing the performance of East Asian and Western Caucasian participants in cognitive and perceptual tasks, challenges the idea of identical thought and visual processes across cultures. Research suggests that the divergence of culture has led to distinct patterns in human cognition and perception.

Previous research has suggested that the cognitive and perceptual differences between East Asians and Western Caucasians are strongly influenced by the social environment and social practice of their respective cultures. For instance, a collectivist culture that places a strong emphasis on the relative positions of individuals within the group and views individuals as constituent parts of social groups may require a set of cognitive mechanisms that cause one to be more sensitive to the relationships between objects and the wider context of scenes. Conversely, an individualist culture, which places stronger emphasis on the value of individuals as separate entities may engender cognitions that are more sensitive to the individual objects within a particular scene (Nisbett & Miyamoto, 2005). Hence, East Asians who are more concerned with people's place as part of a whole may adopt a holistic approach by integrating salient objects and background context while Western Caucasians may employ an analytical style by being more attentive to specific goals (Ji, Peng, & Nisbett, 2000).

Social practices, such as child rearing, may also affect the attentional patterns that participants employ in perceptual tasks like scene or face perception. Fernald and Morikawa (1993) observed mother-and-infant interaction when playing with toys at home. It was found that during play, American mothers tended to label objects more frequently (e.g. "That's a car. It's got nice wheels") than Japanese mothers who used objects to engage their child in social routines (e.g. "Here! It's a vroom vroom. I give it to you. Now give this to me"). Directing a child's attention to certain aspects of play (e.g. the label and attributes of a focal object versus the relationship between an object and its context) may foster different attentional patterns in East Asians and Western Caucasians.

Attentional differences between the two cultures can also be found in maternal speech and toddlers' use of language (Tardif, Gelman, & Xu, 1999). Results show that Mandarin-speaking mothers utilize more verbs than nouns while English-speaking mothers produce more nouns than verbs when communicating with their toddlers. Toddlers showed the same language differences when the proportion of nouns relative to verbs was assessed. These findings imply that communication practices may lead a child to focus either analytically on the object (noun) or holistically on the relationship between the object and context (verb).

In the following sections, the cognitive and perceptual differences observed in participants from different cultures will be discussed in more detail.

2.1.1. Cultural Differences in Cognition

"Correspondence bias," which describes the tendency to attribute behavior to an individual's disposition rather than situational factors, is a well-known phenomenon in social psychology. In an early study conducted by Jones and Harris

(1967), participants were asked to read an essay that took a positive or negative view on a social issue. Although the participants were told that the writer had been required to take a stand for or against the issue, participants tended to ignore the situational constraints and interpreted the content of the essay as the writer's genuine attitude towards the issue. However, cross-cultural studies that manipulated the salience of situational constraints have found that East Asian observers are more likely than Western Caucasians to explain (Choi & Nisbett, 1998) and predict (Norenzayan, Choi, & Nisbett, 2002) behaviors in terms of situational factors.

As an extension of Jones and Harris' (1967) study, Choi and Nisbett (1998) manipulated the salience of situational constraints by exposing participants to the same procedure as the writers (i.e. being required to write an essay either supporting or opposing an issue regardless of their personal opinion). When participants had gone through the same procedure as the writer, participants' correspondence bias was measured by correlating participants' responses to two questions: 1) "How much did the target person express his genuine attitude toward capital punishment in his essay?" and 2) "How much did you express your genuine attitude toward capital punishment in your essay?" Results showed that American participants tended to report that the writer's essay reflected his genuine attitude more so than their own essays reflected their attitude, indicating that exposure to the same experience as the writer did not reduce American participants' correspondence bias when attempting to explain the writer's opinion. Conversely, Korean participants were able to incorporate their experiences (of having gone through the same situation as the writer) to infer the writer's opinion, resulting in a reduced correspondence bias.

Cultural differences are not only present in participants' explanations of behaviors but also their predictions. In Norenzayan et al. (2002)'s study, when

participants were asked to predict a character's future behaviors based upon a given trait (e.g. helpfulness) or narrations illustrating concrete behaviors that reflect the trait, Korean and American participants were equally susceptible to the correspondence bias. However, when situational information was manipulated, Korean participants corrected their correspondent inferences when situational information was more salient, while the manipulation had no effect on Americans. To further investigate the perspective of individuals from different cultures on personality, participants were asked to rate how strongly they agree with multiple statements representing the entity theory of personality (i.e. everyone has a certain personality and it is something that they can't do much about). Results show that Koreans tend to believe that personality may differ according to situational circumstances whereas Americans were more likely to believe that personality traits do not vary by situation.

2.1.2. Cultural Differences in Visual Perception

Cultural differences have been found not only in the way people conceptualize information but also how they perceive visual fields. Studies have shown that East Asian participants are more susceptible to visual illusory tasks because of their dependence on contextual information. Ji, Peng, and Nisbett (2000) examined the strength of the rod and frame illusion by utilizing the classic machine in which a 16 inch square frame can be rotated independently of the rod that is placed within the frame. When asked to judge or rotate a rod until it appears to be vertical (i.e. straight up and down) regardless of the position of the frame, Western Caucasians made fewer errors and spent less time on the task, suggesting that they employed an analytical perceptual style by focusing on the rod relatively independently of the frame. In contrast, the dependence on contextual information made the task more difficult for

East Asians as they were affected by the angle of the frame. Focusing on the focal object (i.e. the rod) and ignoring the context (i.e. the frame) enabled Western Caucasians to perform better than East Asian observers.

Similarly, Doherty, Tsuji, and Phillips (2008) found comparable results in the way Japanese and British participants perceive sizes. In a visual size perception task similar to Ebbinghaus illusion, participants were required to identify which of the two center circles looked larger in two different 3 x 3 arrays. In this illusion, size contrast may impair size discrimination; hence, the center circle may appear larger if it is surrounded by smaller circles as opposed to larger ones. Results showed that British participants were more accurate in discriminating the sizes while Japanese participants were prone to making more errors. This indicates that Japanese participants' integration of the focal object with its context impaired their accuracy in size discrimination.

Studies on scene perception also found similar patterns in the way East Asian and Western Caucasian participants perceive visual scenes. In one study, Miyamoto, Nisbett, and Masuda (2006) assessed participants' ability to detect the changes between two photographs after being primed with Japanese and American scenes. Prior to the main experiment, preliminary assessment of the photographs of small, medium, and large cities in Japan and the United States was carried out. Objective and subjective measures were taken by analyzing the number of objects present in each image and participants' ratings of the complexity and ambiguity of each image (e.g. How ambiguous is the boundary of each object? To what degree do there seem to be parts of the scene that are invisible? To what degree is the scene either chaotic or organized?). Results showed that Japanese sceneries, in general, were more complex and ambiguous than American ones; moreover, the complexity of the scenes was not

different within the three sizes of cities in Japan, but big cities in the United States were significantly more complex than smaller ones. Because the Japanese perceptual environment is far more complex and ambiguous than in the United States, people living in Japan may perceive visual information more holistically by attending to contextual information more often than people living in the United States. Results showed that Japanese participants were able to detect more contextual changes than American participants, suggesting that Japanese observers tend to focus more on the background while American observers directed their attention to the focal objects. These findings also imply that the perceptual style may be driven by the environment an observer is exposed to.

Another study by Masuda and Nisbett (2001) assessed participants' recognition accuracy in scene perception. Participants were shown a series of scenes with focal objects. When prompted to describe the scenes, Japanese participants tended to provide more information about the background than American observers. When the focal objects were paired with the original or novel backgrounds during recognition, Japanese participants were able to recognize the focal objects more accurately when they were shown with the original background as compared with the novel ones; however, the manipulation of old and new background did not have an effect on American observers. This result implies that Japanese participants may encode information holistically by integrating the focal object with the original background that it was presented with.

A further study examining eye movement strategies during scene perception found that American participants tended to fixate quicker and longer at focal objects whereas Chinese participants made more saccades to the background and fixated on the focal object and background for comparable length of time (Chua, Boland, &

Nisbett, 2005). As a result of American participants' attentiveness to the foreground, they were more accurate in recognizing the focal objects even when presented with a new background. Cultural differences found in the eye movement strategies that participants employ provide insight as to how observers allocate attention and encode visual information when perceiving visual fields.

2.1.3. Summary

Although many aspects of human cognition and perception have long been assumed to be identical across cultures, cross-cultural studies provide evidence that human thought and visual processes are heavily influenced by the social environment and social practice of individual cultures. The Eastern collectivist culture, with its focus on the relationships between individuals within groups, may have led East Asians to assimilate focal objects with background context whereas the Western individualist culture, with its emphasis on the value of individuals as independent entities, may lead Western Caucasians to be more analytical by focusing on the salient objects more than contextual information. These two distinct perceptual styles have been identified in various visual processing tasks, such as susceptibility to visual illusions and scene perception; East Asian observers have a tendency to pay more attention to context while Western Caucasians focus more frequently on the focal object. The contrasting holistic-analytic styles associated with Eastern and Western cultures are reflected in the varying thought and visual processes in both cultures and may possibly extend to the eye movement strategies that East Asian and Western Caucasian observers employ in face perception tasks, such as face recognition and emotion recognition (Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Jack, Blais, Scheepers, Schyns, & Caldara, 2009).

2.2. Cultural Differences in Face Perception

Recent face processing studies, such as face recognition and emotion recognition, have revealed systematic cross-cultural differences in the eye movement strategies employed by East Asian and Western Caucasian observers (Kelly, Liu, Rodger, Miellet, Ge, & Caldara, 2011b; Blais et al., 2008; Jack et al., 2009). In a face recognition study by Blais et al. (2008), East Asian and Western Caucasian participants learned a series of own- and other-race faces displaying neutral, happy or disgust expressions. After a brief pause, participants were required to determine whether each of another set of faces, containing previously seen and novel faces, had been shown previously by pressing on a two button control pad. Upon completion of the face recognition task, participants were instructed to categorize each face by its respective race. Results showed that across time and tasks, East Asian participants focused more on the center of the face (in the nose area) indicating a holistic style of processing. Western Caucasians, on the other hand, made more fixations distributed between the eyes and mouth (in a triangular looking pattern) which, they argued, suggests a more analytical processing of individual features. The preference for a holistic or analytical perceptual strategy used for face recognition is in line with the strategy used for other cognitive and perceptual tasks for each racial group (e.g. Choi & Nisbett, 1998; Norenzayan, Choi, & Nisbett, 2002), and was replicated in a younger population when performing a face recognition task (Kelly et al., 2011b).

British and Chinese children aged seven to twelve years were asked to learn and subsequently recognize twelve East Asians and twelve Western Caucasians posing neutral facial expressions in two separate blocks. Eye tracking results showed that British children fixated more frequently on the eyes and mouth while Chinese children made more fixations on the nose region (Kelly et al., 2011b). The fixation

pattern employed by East Asian and Western Caucasian children is similar to those employed by their same-race adults (Blais et al., 2008).

In an emotion recognition study by Jack et al. (2009), East Asian and Western Caucasian participants were required to categorize own- and other-race faces displaying Ekman's six universal emotions ("happy," "surprise," "fear," "disgust," "anger," and "sad"; 1970) plus "neutral." Behavioral results showed that East Asian observers had difficulty recognizing "fear" and "disgust," which were consistently confused with "surprise" and "anger," while Western Caucasian observers had relatively high accuracy at recognizing all facial expressions. Further analyses of eye movements revealed that East Asian and Western Caucasian observers adopt different fixation patterns to categorize facial expressions of emotion. East Asians tended to recognize emotion by focusing on the eye region alone while Western Caucasians looked evenly across the face by focusing on the eyes and mouth. The lack of focus on the mouth may have contributed to East Asian observers' misidentification of relevant information and led to the miscategorization of certain emotions, such as "fear" and "disgust." When in doubt of their answers, East Asian observers tended to choose a less threatening emotion, such as "surprise." These results imply that certain emotions may be expressed differently in the East as observers from this culture have learned to adopt a different fixation pattern to categorize facial expressions.

A recent eye tracking study aimed to examine the influence of contextual expressions on American and Chinese participants' emotion recognition performance and eye movement strategy (Stanley, Zhang, Fun, & Isaacowitz, 2013). Participants were shown slides, in which faces of five identities were presented in a cross shape with the target faces surrounded by a red box, and were instructed to give verbal responses as to which emotions the target faces were displaying (i.e. "happy,"

"surprise," "fear," "disgust," "anger," or "sad"). Slides were either congruent (i.e. all five faces displayed the same emotion) or incongruent (i.e. the four contextual faces displayed the same emotion while the target face displayed a different emotion). Results showed that American participants recognized facial expressions of emotion better than Chinese participants. Analyses on participants' target preference score (which reflected the extent to which participants preferred fixating on the target face over contextual faces) revealed that American participants tended to fixate longer on target faces than contextual faces as compared to Chinese participants, although this was only true for "happy," "surprise," and "anger" faces. The task in this study favored the use of a contrasting strategy (by differentiating the target face from the contextual faces) instead of an integrating strategy to recognize emotions. Although both American and Chinese participants exhibited the contrasting strategy (as nearly half of the transitions between fixations were made between target and contextual faces, rather than between different contextual faces), the percentage of target transitions was greater for American than Chinese participants. This suggests that Americans showed a greater tendency to utilize the contrasting strategy, although this effect was only significant for "happy," "surprise," "fear," and "sad" faces. Additionally, the contrasting strategy was found to be correlated with higher recognition accuracy of "happy," surprise," and "sad" faces. These findings suggest that although both American and Chinese participants were affected by contextual faces when recognizing emotions, employing the contrasting strategy by moving between the target and contextual face had a positive impact on American participants' recognition accuracy while focusing on contextual faces with fewer target transitions was not as beneficial for Chinese participants.

Eye tracking results from these face processing studies not only revealed systematic differences in the eye movement strategies employed, but also suggest an adaptation of strategies according to the nature of the task. More specifically, East Asian observers tended to focus in the nose area when performing a face recognition task but shifted their attention towards the eyes when asked to identify emotion rather than identity (Blais et al., 2008; Jack et al., 2009). This change of eye movement strategy suggests that certain emotions may be expressed differently in the East, leading observers to have learned to adopt a different fixation pattern to recognize emotions. Behavioral results showed that East Asian observers' lack of focus on the mouth may have contributed to the misrecognition of certain emotions such as "fear" and "disgust," which were consistently confused with "surprise" and "anger" (Jack et al., 2009). On the other hand, Western Caucasian observers appear to employ a similar strategy for both face recognition and emotion recognition tasks by moving between the eyes and mouth to recognize identity and emotion (Blais et al., 2008; Jack et al., 2009). Western Caucasian observers may not have altered their fixation pattern in the emotion recognition task possibly because fixating on the expressive regions of the face (i.e. the eyes and mouth) was an optimal strategy to recognize emotion as well as to recognize identities. Behavioral results showed that the even distribution of fixations across faces enabled Western Caucasians to recognize all facial expressions of emotion with relatively high accuracy.

Although cross-cultural studies have identified distinct Eastern and Western eye movement strategies employed in face recognition and emotion recognition tasks, studies investigating the role of genetic and cultural factors revealed that Asian participants who are more familiar with the Western culture (e.g. British born Chinese; Kelly, Jack, Miellet, De Luca, Foreman, & Caldara, 2011a) tended to adopt either an

Eastern or Western strategy, suggesting that there may be more to cross-cultural differences than a simple Eastern-Western distinction (see Chapter Three for a more detailed discussion). Culture alone may not fully explain the diversity in observers' eye movement strategies; exposure to and familiarity with the diagnostic features of other-race faces encountered during development may also affect one's looking pattern. Therefore, investigating how observers from unique cultures, such as Malaysia, would perform in face perception tasks will give us insight as to whether observers originating from Eastern countries that are greatly influenced by the Western culture would employ eye movement strategies more alike with East Asians or more in line with those of British born Chinese.

2.3. The Malaysian Culture

The majority of cross-cultural studies conducted so far have employed a distinction between East Asian and Western Caucasian cultures and comparisons have tended to be made between countries with relatively low degrees of multiculturalism; more specifically, China is composed of 91.9% Han Chinese and 8.1% of Zhuang, Uygur, Hui, Yi, Tibetan, Miao, Manchu, Mongol, Buyi, Korean, and other nationalities, Japan is composed of 99% Japanese and 1% of Korean, Chinese, Brazillian, Filipino, and other, while the United Kingdom is composed of 83.6% English, 8.6% Scottish, 4.9%Welsh, 2.9% Northern Irish, 2% black, 1.8% Indian, 1.3% Pakistani, 1.2% mixed, and 1.6% other ("Ethnicity and Race by Countries," 2005). This has enabled researchers to obtain groups as distinctive as possible from one another for comparison. However, there are many countries within both the West and East Asia which do not fall so easily into this dichotomy. Although Malaysia is an East Asian country, it is strongly multicultural and heavily influenced by the Western

culture. Malaysia was colonized by the British in the early 19th century. During this period, the British exerted Western influence on the Malaysian culture by making systematic changes to the structure of the country, including its education system (Advameg, 2011).

Before the European colonial powers were established, the Malaysian school system comprised religious classes conducted by Muslim missionaries, which later transformed into formal religious schools. Although British influence on education can be traced back to early 19th century, it was not until 1867 when the Malay States were transferred to the Colonial Office in London that the British started placing more emphasis on the education system by taking measures such as appointing an inspector of schools, reforming the vernacular education, and increasing grants to all English medium schools. The Education Code implemented in 1899 laid the foundation of Malaysia's education system and emphasized the importance of learning English by making it one of the core subjects to be taught in all schools. Even though the four major education systems in Malaysia share the same basis, the Malay, Chinese, and Tamil schools continue to uphold their cultural identity by promoting their respective worldview while the English medium schools encouraged the acquisition of Western values. The English schools, largely missionary or governmental, welcomed students from all ethnic backgrounds and created an environment that encouraged respect for one another as individuals rather than as members of certain ethnicity. It also fostered a common bond across ethnic groups—a mutual Western-orientated outlook (Gaudart, 1987).

Although Malaysia gained its independence in 1957, Western influence can still be traced in the Malaysian culture. For instance, Malaysian FM radio stations are composed of 40% Malay stations, 26% English, and 15% Chinese (Davies, 2011). 86%

of movies shown in Malaysian cinemas are Western and a mere 14% are local. This compares to 57% local movies in Chinese and Japanese cinemas (Epstein, 2011).

The population of Malaysia is also highly diverse. In the early years, Malaysia, a promising land with vast reserves of tins and fertile soil for rubber trees, has attracted immigrants from China and India to come and work as laborers while the Malays worked in small holdings and rice cultivation ("Background Note: Malaysia," 2011). Until today, the ethnic composition of Malaysia remains highly diverse, with Malays composing 50.4% of the population, Chinese 23.7%, indigenous 11%, Indian 7.1%, and others composing 7.8% (including Western Caucasians and Africans; "The World Factbook," 2011). Data from the Department of Immigration in 2006 reported that 8.6% of the country's 19,444 expatriates originated from the United Kingdom, placing it as the fourth largest source of expatriates to Malaysia. Conversely, no African (or black majority) country was listed as a major source of immigrants in any class of international migration (Kanapathy, 2008). Major festivals from each ethnic background (e.g. Hari Raya for Malays, Chinese New Year for Chinese, and Deepavali for Indians) continue to be celebrated by all people in Malaysia on a national level.

2.4. Aims of This Thesis

Previous cross-cultural studies have identified a distinct holistic-analytic pattern that observers employ in various cognitive and perceptual tasks. Recent face perception studies utilizing eye tracking methodologies have also revealed distinct Eastern and Western viewing patterns when recognizing identities and facial expressions of emotion (Blais et al., 2008; Jack et al., 2009). However, studies exploring genetic and cultural factors found that British born Chinese observers

employed either Eastern or Western eye movement strategies, suggesting that a simple Eastern-Western distinction does not fully explain the diversity in observers' eye movement strategies (Kelly et al., 2011a; see Chapter Three for an examination of the data). This thesis aimed to investigate Malaysian Chinese participants' eye movement strategy and recognition abilities by requiring participants to perform static and dynamic face recognition (Chapter 3 and Chapter 5) and emotion recognition (Chapter 4 and Chapter 6) tasks on African, East Asian, and Western Caucasian faces.

2.4.1. Methodological Considerations

As discussed in Chapter One, most face processing studies utilize static stimuli that do not accurately represent a natural social situation. Faces are complex in nature; however, perceiving photos of faces in a highly controlled and artificial setting may reduce the complexity of face perception tasks.

In face recognition, for example, an eyewitness who witnessed a crime is usually taken by surprise and may be required to process vast visual information (e.g. the suspect, weapons, chaotic scene) in a limited amount of time; conversely, participants taking part in a laboratory experiment are given a standardized amount of time to view the target face without other visual distractions and may even have clear instructions of what the task entails (e.g. you will be asked to recall these faces during the recognition phase). Moreover, an eyewitness does not usually have a clear view of the target face that may have been intentionally disguised (e.g. masks or sunglasses) and the eyewitness will be required to identify the culprit from a different point of view with possible natural changes in appearance (e.g. aging or facial hair), whereas participants are normally required to recognize the target face from an identical photograph. Studies examining factors such as intentional disguises, natural changes

in appearance, or rotation of faces in depths have found that these manipulations have a negative impact on participants' recognition abilities, indicating that a change of perspective and/or presentation method, which would naturally occur when extending laboratory studies to real life situations, would result in poorer recognition performance (Hockley, Hemsworth, & Consoli, 1999; Read, Vokey, & Hammersley, 1990; Valentin, 1999; Hill, Schyns, & Akamatsu, 1997; O'Toole, Jiang, Roark, & Abdi, 2006).

For facial expressions of emotion, natural emotion tends to unfold gradually and subtly while photographs merely capture a single intense emotion which may be posed and exaggerated. Identifying one's emotion during a typical social interaction is more complex than identifying a posed emotion portrayed in a photograph because of numerous factors such as varying display rules which govern the appropriateness of one's display of emotion in different cultures (e.g. de-intensifying, neutralizing, or masking one's actual emotion; Ekman, 1970). Thus, as compared to the complex nature of the task in a natural social situation, presenting exaggerated full-blown facial expressions of emotion in laboratory studies would result in better recognition performance.

In the current thesis, Chapter Three and Chapter Four focused on the examination of Malaysian Chinese participants' eye movement strategy and sensitivity in recognizing identities and emotions respectively when perceiving photos of faces. The behavioral and eye tracking results from these two studies would be comparable to previous eye tracking studies involving different populations (e.g. East Asians, Western Caucasians, and British born Chinese), providing insight as to whether Malaysian Chinese employed a strategy more alike with East Asians or more in line with those of British born Chinese. However, because of the methodological

considerations regarding static stimuli, naturalistic videos of models (term used as short-hand—the volunteers were not trained models) introducing themselves and describing different emotional experiences were collected for Chapter Five and Chapter Six respectively to investigate whether Malaysian Chinese participants adopt different eye movement strategies when viewing a different type of stimuli which more accurately represent a natural social situation. Additionally, participants' behavioral performance in recognizing dynamic identities and emotions will also be examined.

Chapter Three

Static Face Recognition

(A version of this chapter was published as Tan, C. B. Y., Stephen, I. D., Whitehead, R., & Shepard, E. (2012). You look familiar: How Malaysian Chinese recognize faces. PLoS ONE 7 (1):e29714. doi:10.1371/journal.pone.0029713.)

Abstract

East Asian and Western Caucasian observers employ different eye movement strategies for a variety of visual processing tasks, including face processing. Recent eye tracking studies on face recognition found that East Asians tended to integrate information holistically by focusing on the nose while Western Caucasians perceive faces analytically by moving between the eyes and mouth. The current study examined the eye movement strategy that Malaysian Chinese participants employed when recognizing African, East Asian, and Western Caucasian faces. Rather than adopting the Eastern or Western fixation pattern, Malaysian Chinese participants use a mixed strategy by fixating most on the nose, followed by the eyes, then the mouth. The combination of Eastern and Western strategies proved advantageous in participants' ability to recognize East Asian and Western Caucasian faces, suggesting that individuals learn to use eye movement strategies that are optimized for recognizing the faces with which they are more familiar.

3.1. Introduction

Human faces bear a large amount of information about individuals' identities and characteristics (including gender, age, health, and attractiveness). Although

human faces are similar, in that they comprise the same basic features in approximately the same configuration, they are easily distinguished by observers. Evidence from neuroimaging studies suggests that there is a specialized mental module, possibly located in the fusiform face area (Kanwisher & Yovel, 2006) dedicated to face processing. This ability to recognize different faces proficiently may have social and evolutionary advantages, including allowing us to remember specific individuals' behavior in social situations and recognizing cooperators and defectors (Volstorf, Rieskamp, & Stevens, 2001; Mealey, Daood, & Krage, 1996). Therefore, investigating how people extract visual information in relation to their ability to recognize faces is of particular interest for human social behavior (Peterson & Rhodes, 2003).

3.1.1. Other-Race Effect

The other-race effect (also known as the own-race bias or the other-race disadvantage) is the phenomenon in which people's ability to recognize faces from their own race is more accurate than for faces from other races. Meta-analytic reviews revealed that the other-race effect is a reliable phenomenon across various experimental conditions, with 80% of the samples showing a significant other-race effect (Meissner & Brigham, 2001; Bothwell, Brigham, & Malpass, 1989).

Developmental studies have revealed that the other-race effect emerges in Caucasian and Chinese infants as young as six months old (Kelly, Quinn, Slater, Lee, Ge, & Pascalis, 2007; Kelly, Liu, Lee, Quinn, Pascalis, Slater, & Ge, 2009). Both studies, using an identical procedure, comprised of two phases: the habituation phase and the test phase. During the habituation phase, infants viewed images of faces on a screen while their eye movements were recorded on a camera. The experimenter

would press a key when the infant first fixated on the image and release the key when the infant looked away. When the habituation criterion was met (i.e. the infant fixated on the images equal to or less than 50% of the average fixation duration from the infant's first two trials), the infant would proceed to the test phase, which consisted of two trials. The infant was first presented with two images, and the experimenter would press on a key to start a five seconds countdown when the infant first looked at the images. After five seconds, the images disappeared and reappeared again with their left or right position reversed. The experimenter would start another five second countdown when the infant fixated on the images. Eye movements that were recorded throughout the experiment were analyzed and coded by two observers, and recognition was determined by the infant's preference for a new face (i.e. longer fixation duration) across the two trials; in other words, shorter fixation duration would indicate recognition of a face as infants have shorter attention span for stimulus previously seen. Results showed that three-month-old Caucasian infants were able to recognize all four included races of faces (i.e. Caucasian, African, Middle Eastern, and Chinese). Six-month-old infants showed an emergence of the other-race effect by demonstrating recognition of Caucasian and Chinese faces only. Caucasian infants demonstrated a significant other-race effect at the age of nine months, when only Caucasian faces were reliably recognized (Kelly et al., 2007). Similarly, a study with Chinese infants revealed that three-month-olds were able to recognize all three included races of faces (i.e. Chinese, African, and Caucasian), showing no other-race effect. However, by the age of six months, Chinese infants were able to recognize own-race faces and marginal recognition for Caucasian faces, displaying an emergence of the other-race effect. At nine months, infants' recognition was best for

Chinese faces only (Kelly et al., 2009). These findings suggest that infants' ability to recognize own-race faces improved substantially as compared to other-race faces.

Research examining genetic and cultural factors showed that the cognitive mechanisms involved in face recognition show plasticity during childhood after the age of 3, allowing a reversal of the other-race effect (Sangrigoli, Pallier, Argenti, Ventureyra, & Schonen, 2005). Three groups of adults (i.e. Caucasians living in France since birth, native Koreans brought up in Korea, and adults of Korean origin adopted by Caucasian families in France between ages 3 and 9) were presented with a target face at the center of a screen followed by two faces shown next to each other. Both faces remained on screen until participants decided which photo matched the target face by making a key press. Results showed that adopted Korean participants who had more experience with Caucasian faces were better at recognizing Caucasian than Asian faces, a performance similar to those of Caucasian participants. Conversely, Korean participants recognized Asian faces better than Caucasian ones. Previous studies have found that the other-race effect emerges as early as six months old (Kelly et al., 2007; Kelly et al., 2009), suggesting that Koreans adopted by Caucasian families between ages 3 and 9 would have learned to recognize Asian faces before moving to France. Thus, the findings imply that adopted Koreans may have lost the advantage to recognize Asian faces due to the exposure to Caucasian faces, indicating that the other-race effect can be reversed even in late childhood.

A further study involving 6-month-old Caucasian infants explored the role of perceptual training on the other-race effect (Heron-Delaney, Anzures, Herbert, Quinn, Slater, Tanaka, Lee, & Pascalis, 2011). In pre-training testing, infants fixated longer on Chinese than Caucasian faces and novel than familiar faces, providing evidence that infants show the other-race effect and an ability to recognize faces by the age of 6

months. Parents were then instructed to show their children a book of either Chinese or Caucasian faces for 2 to 3 minutes each day for a week, every other day for the following week, then less frequently according to a fixed schedule provided. Infants would have received a total of 70 minutes of exposure to own- or other-race faces over the course of 3 months. After 3 months of training, infants who were now 9-months-old were required to recognize faces from the book in different viewpoints (i.e. exact image from the book or three-quarter pose of the same face). Results revealed that Caucasian infants who received training on Chinese faces looked longer at novel than familiar Chinese faces whereas infants who received training on Caucasian faces spent similar amount of time looking at novel and familiar Chinese faces. The results indicate that infants were able to discriminate between Chinese faces with training on other-race faces while infants who did not receive other-race training could not discriminate between Chinese faces, replicating the findings of Kelly et al. (2007). These findings suggest that although the cognitive mechanisms involved in face recognition becomes finely tuned to own-race faces early in life, exposure to and familiarity with other-race faces can reduce the other-race effect by allowing observer to develop a face prototype based on the type of facial input or training received.

Another study that trained adults to differentiate African American or Hispanic faces at an individual (e.g. Joe or Bob) or categorical level (e.g. African American or Hispanic) also showed that training can improve one's recognition ability (Tanaka & Pierce, 2009). In a pre-training face recognition test, participants learned 12 African American and 12 Hispanic faces and were required to identify whether the 24 African American and 24 Hispanic faces presented during the recognition phase were "old" or "new." Participants were then trained to individuate

African American faces and categorize Hispanic faces or the opposite. In the learning task of the training, individuation race of faces were paired with unique labels (e.g. A, S, D, F, G, H, J, K) while categorization race of faces were paired with the same label (e.g. O). Participants were then asked to name each face according to the label given in the learning task and had to repeat the block until they were able to recognize all faces correctly. After 7 blocks of learning and naming tasks, participants were required to perform another naming task but under time constraints. Again, participants had to recognize all faces correctly under different time constraints (i.e. 3000 msec for the first block, 2500 msec for the second, and 2000 msec for the third block). Finally, participants had to perform a post-training face recognition task, with an identical procedure and face stimuli used in the pre-training test. Results showed that participants' recognition performance improved after both types of training (i.e. individuation or categorization); moreover, participants who received the individuation training improved more than participants who received the categorization training. The findings of this study provide support to the expertise hypothesis, indicating that sufficient training at a subordinate (or individual) level allowed observers to recognize other-race faces better by encoding features of the face, whereas training at a basic level of race shifted observers' attention to race- rather than individual-specific features.

Studies involving trivia questions as a distractor task in between the learning and recognition phases found that the other-race effect arises as a result of retention or retrieval deficits, not encoding-based effects (Papesh & Goldinger, 2009). In this series of experiments, participants were presented with a target face, followed by an interstimulus interval before they were asked to discriminate a target face from a manipulated version of itself (e.g. distorted in size of features or perceived lightness

of face). Results showed that the manipulation of retention intervals and stimuli did not result in a strong other-race effect. The other-race effect had an impact on participants' recognition accuracy only when a distractor was presented. Caucasian participants were able to recognize own-race faces without difficulty even when presented with trivia questions during the retention interval; however, Asian faces were more susceptible to distraction, resulting in poorer recognition accuracy (Papesh & Goldinger, 2009).

Cross-cultural studies investigating Caucasian and African observers' verbal descriptions of own- and other-race faces found that people from different racial groups use different facial features to describe faces (Ellis, Deregowski, & Shepherd, 1975). Caucasians, for instance, were more likely to describe features like the eyes, while Africans focused on the lower facial features such as the nose and lips. In conjunction with this finding, Hills and Lewis (2006) hypothesized that training Caucasian participants to focus on the critical features of African faces will subsequently increase their recognition accuracy. Results showed that the other-race effect was significantly reduced after participants received feature-critical training, implying that shifting one's attention to the diagnostic features relevant to a specific race of face improves participants' recognition accuracy. This suggests that the other-race effect may be attributable to the use of looking strategies, learned during development, that are optimized for recognizing familiar own-race faces, but which are suboptimal when applied to unfamiliar other-race faces.

Although the other-race effect is a robust phenomenon sustained by meta-analytic studies (Meissner & Brigham, 2001; Bothwell et al., 1989), only recently have researchers questioned whether the same underlying attentional strategies are deployed when observers of different races perceive faces.

3.1.2. Eye Tracking

Systematic cross-cultural differences have been found in the fixation patterns of East Asian and Western Caucasian children and adults (Kelly, Liu, Rodger, Miellet, Ge, & Caldara, 2011b; Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Jack, Blais, Scheepers, Schyns, & Caldara, 2009). In the learning phase of a face recognition study, East Asian and Western Caucasian participants were shown a series of own- and other-race faces displaying neutral, happy or disgust expressions. Participants were then required to decide whether each of another set of faces, which contained previously seen and novel faces, had been shown in the learning phase by pressing on a two button control pad. Upon completion, participants categorized each face by its respective race. Across time and tasks, East Asian participants fixated more on the center of the face (in the nose area) which indicate a holistic style of processing. On the other hand, Western Caucasians made more fixations distributed between the eyes and mouth (in a triangular looking pattern), suggesting a more analytical processing of individual features. Observers' preference for a holistic or analytical perceptual strategy used for face recognition is in line with the strategy used for other cognitive and perceptual tasks for each racial group (e.g. Choi & Nisbett, 1998; Norenzayan, Choi, & Nisbett, 2002), and was replicated in a younger population when performing a similar face recognition task (Kelly et al., 2011b).

Chinese and British children, age ranging from seven to twelve years, were required to learn and subsequently recognize twelve East Asians and twelve Western Caucasians posing neutral facial expressions in two separate blocks. Eye tracking results revealed that Chinese children made more fixations on the nose region while British children fixated more frequently on the eyes and mouth (Kelly et al., 2011b).

The fixation pattern identified in Chinese and British children were similar to those observed in East Asian and Western Caucasian adults (Blais et al., 2008).

Research revealed that observers from different cultures employ different fixation patterns when perceiving faces. However, the underlying mechanism in face processing remains unidentified. In a study by Caldara, Zhou, and Miellet (2010), researchers aimed to identify the facial information extracted by East Asian and Western Caucasian observers by using the *Spotlight* technique to restrict the facial information available to observers in varying degrees. Results revealed that observers from both cultures use identical facial information (i.e. the eyes and mouth) to recognize faces when facial information outside of the central vision was restricted within 2 and 5 degrees. However, when visual information from the eyes and mouth is readily available in the 8 degrees condition, East Asian observers employed a different fixation pattern by focusing on the center of the face, in the nose area. This finding suggests that the same facial information is utilized to recognize faces although East Asian and Western Caucasian observers employ different fixation patterns to achieve recognition.

A further study investigating the role of genetic and cultural factors found that British born Chinese performed equally well at recognizing East Asian and Western Caucasian faces. When recognizing own- and other-race faces, British born Chinese employed either an Eastern or Western eye movement strategy, with fixations landing predominantly around either the eyes or nose region (Kelly, Jack, Miellet, De Luca, Foreman, & Caldara, 2011a). The authors suggested that the utilization of either an Eastern or Western strategy exemplified an influence of both cultures.

However, our examination of the data (Tan, Stephen, Whitehead, & Sheppard, 2012) suggests that most participants actually employed intermediate looking strategies, with none of the 20 participants showing a bias of over 10% towards either strategy. Further, the bias values are normally distributed (Kolmogorov-Smirnov test p=0.2) around a mean of 2.5%, not bimodally distributed as would be expected if participants could be validly classified into either Western or Eastern strategists (Tan et al., 2012), see Figure 3.1 for a visual representation of the data. This finding suggests that although British born Chinese participants are genetically Asian, exposure to and familiarity with the Western culture and/or faces fostered an intermediate eye movement strategy that is advantageous for recognizing both own- and familiar other-race faces.

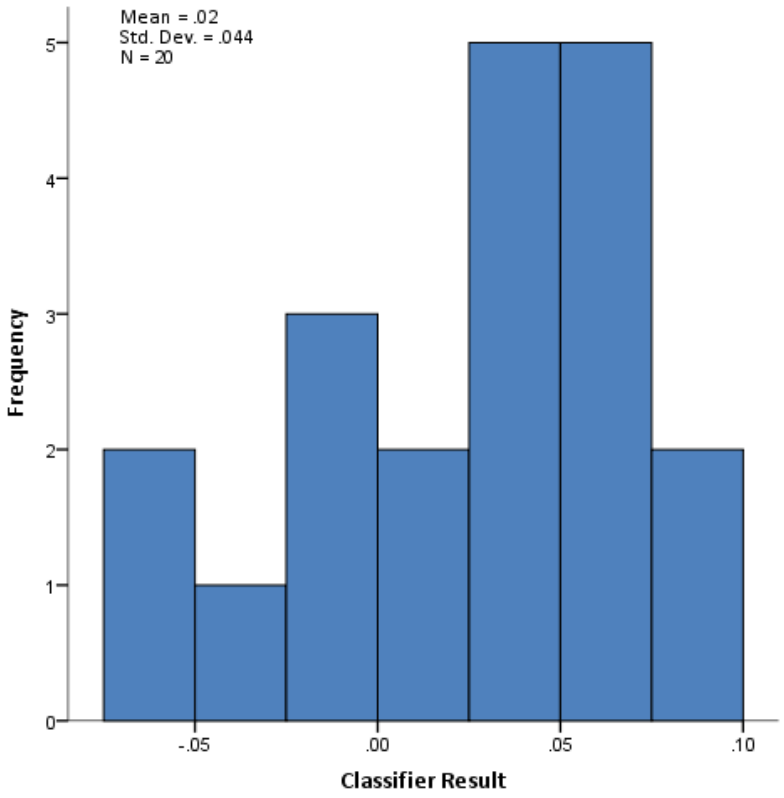


Figure 3.1. Histogram of British born Chinese participants' classifier result. Lower negative values indicate a stronger bias to employ a Western strategy (by fixating more on the eyes) whereas higher positive values indicate a greater tendency to employ an Eastern strategy (by fixating more on the nose).

3.2. Current Study

The current study aimed to investigate Malaysian Chinese participants' recognition sensitivity and eye movement strategy when performing a face recognition task involving own- and other-race faces (i.e. African, East Asian, and Western Caucasian faces). Although eye tracking studies have revealed distinct fixation patterns that Eastern and Western observers employ (Blais et al., 2008), studies exploring genetic and cultural factors found that British born Chinese observers utilized a combination of both strategies, implying that exposure to and familiarity with the diagnostic features of other-race faces encountered during development may also affect one's looking pattern (Kelly et al., 2011a, Tan et al., 2012). Thus, examining how Malaysian Chinese observers would perform in a face recognition task will give us insight as to whether participants originating from a unique culture like Malaysia would employ a strategy more alike with East Asians or more in line with those of British born Chinese.

Malaysia is geographically located in the East; therefore, Malaysian Chinese observers may, perhaps, employ a strategy similar to those of East Asians, with more fixations landing on the nose than the eyes and mouth. The lack of attention to the diagnostic features of Western Caucasian and African faces (Ellis et al., 1975; Hills & Lewis, 2006) may result in poorer recognition accuracy for Western Caucasian and African faces, reflecting a strong other-race effect for East Asian faces. On the other

hand, because the Malaysian culture is substantially different from East Asian culture, the influence of the west may be reflected in Malaysian Chinese observers' performance. Previous face recognition study revealed that Asian participants who are more familiar with the Western culture (e.g. British born Chinese) employed a combination of both Eastern and Western strategies, allowing participants to recognize both own- and other-race faces equally well (Kelly et al., 2011a). As presented in Chapter Two, Malaysia is a strongly multicultural country with over 200 years of history with the West (Advameg, 2011). Today, Western influence can still be traced in Malaysia's education system (Gaudart, 1987) and mass media (Davies, 2011; Epstein, 2011). Given the diverse nature of the country, Malaysian Chinese students enrolled in a branch campus of a British university may use a mixture of East Asian and Western Caucasian eye movement strategies to perceive faces, with more fixations landing on the eyes and nose than the mouth. The hybrid of Eastern and Western strategies may enable observers to be equally good at recognizing East Asian and Western Caucasian faces, but not the less familiar African faces.

3.3. Methods

3.3.1. Participants

Twenty-two young adults (10 males, 12 females, mean age 21.86 years) participated in this study. All participants were Malaysian Chinese students attending the University of Nottingham Malaysia Campus, and have not lived outside of Malaysia for more than three years. All participants had normal or corrected vision and were given a bar of chocolate for their participation. Written informed consent

was obtained from all participants and the protocol was approved by the University of Nottingham, School of Psychology Ethics Committee.

3.3.2. Materials

Images were taken of 60 "models" (term used as short-hand—the volunteers were not trained models; 10 African males, 10 African females, 10 East Asian males, 10 East Asian females, 10 Western Caucasian males, and 10 Western Caucasian females) holding a neutral expression without skin makeup. Models held a Munsell N5 painted board over the shoulders and were photographed in a lighting booth painted with Munsell N5 neutral gray paint, with a GretagMacbeth Mini ColorChecker color chart in the frame. The booth was illuminated with d65 fluorescent tubes, in high-frequency fixtures to reduce the effects of flicker, with no other lightings present in the room (Verivide, UK). As described by Stephen, Coetzee, Law Smith, & Perrett (2009), photographs were color corrected using a least-squares transform, from an 11-expression polynomial expansion of camera RGB values for 24 ColorChecker patches to the manufacturer-specified CIELab values of the same patches. The East Asian and Western Caucasian stimuli were obtained from a set of images collected at the University of St Andrews, UK while the African stimuli were collected from the University of Pretoria, South Africa.

The images were aligned on the eyes position and cropped around the face using Psychomorph software (Tiddeman & Perrett, 2001) and the background was edited to show a clear gray background by removing the 24 ColorChecker patches and the painted board held over the models' shoulders using Adobe Photoshop CS. The images, which were 270x333 pixels in size, were presented at a distance of approximately 60cm on a Tobii T60 eye tracker (a 17in TFT monitor with a screen

resolution of 1280x1024 pixels). Presentation of the stimuli was controlled by Tobii Studio software, version 2.3.

3.3.3. Procedure

The experiment involved two phases: the learning phase and the face recognition phase. During the learning phase, 30 faces (5 male, 5 female East Asian, 5 male, 5 female Western Caucasian, and 5 male, 5 female African) were shown on a Tobii eye tracker and participants were asked to rate the faces for attractiveness on a seven-point Likert scale (i.e. 1 for "very unattractive" to 7 for "very attractive"). Participants were not informed that they would be performing a subsequent face recognition task. To distract participants from remembering the faces, participants filled out the Autism Quotient questionnaire (which is a 50 item forced-choice self-report questionnaire that assesses autistic traits) as a distractor task. Upon completion of the questionnaire, participants performed a face recognition task in which 60 faces (10 male, 10 female East Asian, 10 male, 10 female Western Caucasian, and 10 male, 10 female African of which half were new faces) were presented. Participants gave a yes or no response to indicate if they had seen the face before. After the face recognition task, participants were asked to categorize the 30 faces presented in the learning phase by race.

On each trial, a central fixation cross was presented for one second followed by a face presented pseudorandomly in one of four quadrants of the computer screen to avoid fixation bias. The face stimulus was presented for 5 seconds in both phases and was followed by a forced-choice question in relation to the task (i.e. "How attractive do you think this face is?" for the learning task and "Have you seen this

picture before?" for the recognition task). Each response was followed by the central fixation cross, which preceded the next face stimulus.

3.3.4. Eye Tracking

Eye movements were recorded with an on-screen remote eye tracking system (Tobii T60), in which an infrared camera is integrated to the lower part of the 17in TFT monitor. The eye tracker performs binocular tracking at a data sampling rate of 60 Hz (approximately every 17 milliseconds) and has high accuracy (0.5°) and drift compensation (less than 0.3°). This non-intrusive eye tracker allows participants to behave naturally by allowing a large freedom of head movement (44x22x30cm) with no visible or moving tracking device that may affect participants. Each task began with a calibration procedure as implemented in the Tobii Studio software to ensure accurate tracking of eye gaze. Once participants' eyes were detected by the infrared camera, which would be shown as two moving dots on the interface dialogue box, a standard nine-point calibration procedure began to determine the direction of participants' eye-gaze. Fixation, which was the main measurement used in this study, is defined by the standard Tobii fixation filter as two or more consecutive samples falling within a 35 pixel radius.

3.3.5. Data Analyses

A-prime (A') values were calculated to determine participants' recognition sensitivity as d-prime (d') values were not normally distributed. A', which is a non-parametric equivalent of d', indexes participants' sensitivity to old and new faces taking into account both hits (i.e. correct detection of an old face) and false alarms (i.e. incorrect identification of a new face as an old face). An A' value closer to 1.0 reflects higher sensitivity, whereas an A' value closer to 0.5 indicates chance

performance (Grier, 1971; Pallier, 2002; Zhang & Mueller, 2005). The formulas for A' are:

If hit > false alarm, A' = 1/2 + { [(hit – fa) * (1 + hit – fa)] / [4 * hit * (1 – fa)] }

If false alarm > hit, A' = 1/2 + { [(fa – hit) * (1 + fa – hit)] / [4 * fa * (1 – hit)] }

Eye tracking data was processed directly from the eye tracker using the Tobii Studio software. The total number of fixations a participant made within the predefined areas (eyes, mouth, and nose) was accumulated using the area of interest (AOI) analysis, which counted the number of fixations falling within the predefined areas, see Figure 3.2. AOIs were drawn around the eyes, nose, and mouth as internal features were found to be useful for the recognition of both familiar and unfamiliar faces (Ellis et al., 1979). To ensure that the current eye tracking result is comparable with the previous face recognition study utilizing similar procedure (Blais et al., 2008), the AOI sizes were reduced to fit tightly around the eyes, nose, and mouth. Eye tracking result for the new set of AOIs revealed the same pattern of results regardless of the size of AOIs. Each participant completed 60 trials for 5 seconds each. To ensure the validity of the eye tracking data, only participants who had an average total fixation count equal to or more than 5 counts across trials during the recognition task were included in the data set. Two participants (1 male) did not meet the criteria and were excluded from the eye tracking analysis.

The Autism Quotient (AQ) questionnaire, which was presented as a distractor task, is a standardized measure widely used in research which uncovers cognitive features that characterize autism in typical individuals who scored highly on the AQ. Pearson's correlation was conducted to examine whether participants' AQ scores

predict behavioral performance. However, there was no correlation between the two factors, suggesting that participants' AQ scores did not predict recognition abilities.

Statistical analyses were performed using SPSS version 21. Linear Mixed Modeling analyses were conducted to test the effect of task (learning and recognition), race of face (African, East Asian, or Western Caucasian) and feature of face (eyes, mouth, or nose) on the four dependent variables: total number of first fixations (i.e. the accumulated number of times that the first fixation made on a face fell on either the eyes, mouth, or nose for all trials), total fixation count (i.e. the accumulated number of fixations landing on the eyes, mouth, and nose for all trials), total fixation duration (i.e. the accumulated fixation duration (in seconds) landing on the eyes, mouth, and nose for all trials), and mean fixation duration (i.e. total fixation duration divided by total fixation count). Participant ID, face ID, task, race of face, feature of face were included as factors. All main effects and the two- and three-way interactions between task, race of face, and feature of face were included in the model (dependent variable = total number of first fixations, total fixation count, total fixation duration and mean fixation duration; fixed factors = task, race of face, and feature of face). Participant ID and face ID were included as subjects to avoid pseudoreplication. Participant ID nested within task and face ID nested within race were included as random factors for mean duration. For total fixation count and total fixation duration, only participant ID nested within task was included as a random factor because the final Hessian matrix was not positive, which casts doubt on the validity of results; however, all results with or without face ID nested within race included as a random factor was identical. No random factors were included for total number of first fixations because of the same reason; again, all results with or without participant ID

nested within task and face ID nested within race included as random factors were identical.

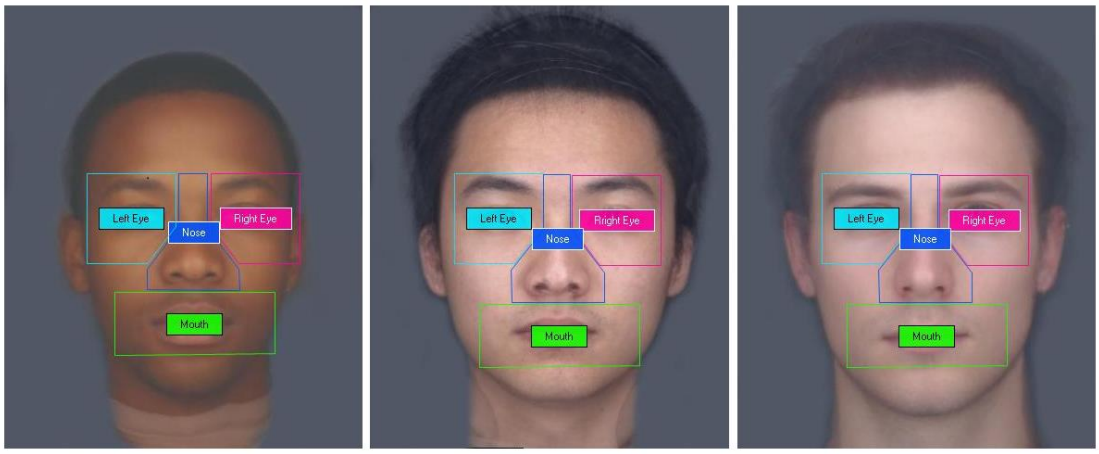


Figure 3.2. The predefined area of interest (AOI) used to analyze eye gaze. Composite images are shown for illustration purposes. Real faces were used in the actual experiment.

3.4. Results

3.4.1. Recognition Sensitivity

One-sample t-tests on A' values showed that Malaysian Chinese participants were able to recognize African, $t(21)=14.85$, $p<.001$, East Asian, $t(21)=32.73$, $p<.001$, and Western Caucasian faces, $t(21)=29.83$, $p<.001$, above chance level (test value = 0.5). A' values were not normally distributed. Hence, data was reflected using a square root transformation (i.e. the square root of one minus A') before analysis. However, the bar graph below (Figure 3.3) represents untransformed A' values for easier interpretation.

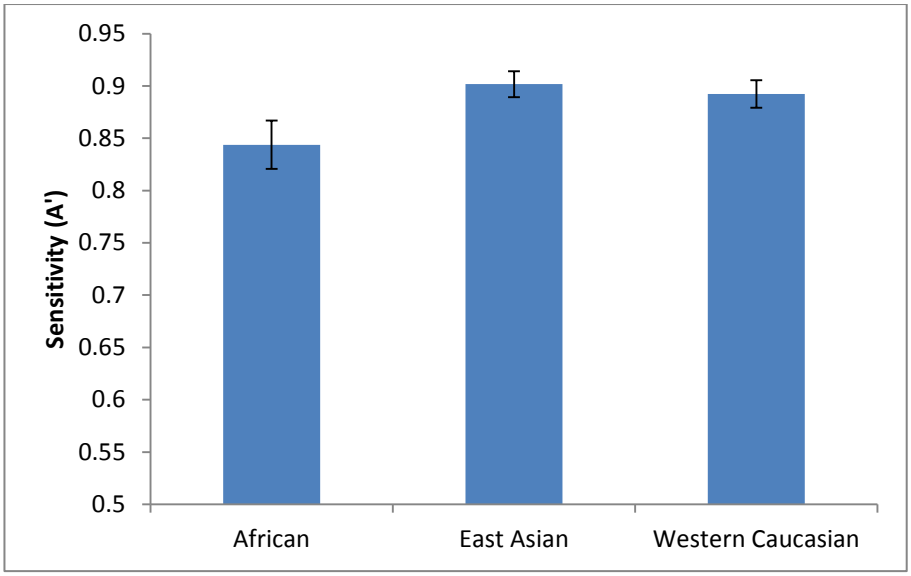


Figure 3.3. Malaysian Chinese participants' recognition sensitivity for African, East Asian, and Western Caucasian faces. Error bars report standard errors of the mean. Participants recognized East Asian faces significantly better than African faces.

A one-way repeated measures analysis of variance (ANOVA) showed that participants' recognition sensitivity (A') differed significantly between races of face, $F(2, 42)=3.96$, $p=.027$. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that participants were significantly better at recognizing East Asian than African faces ($p=.037$). However, recognition sensitivity for Western Caucasian faces was not significantly lower from East Asian faces ($p=1$) or significantly higher than African faces ($p=.226$), see Table 3.1 for participants' hit rates and false alarm rates.

Table 3.1. Hit rates and false alarm rates in recognizing African, East Asian, and Western Caucasian faces.

	Hit Rates (SD)	**False Alarm Rates (SD)**
African faces	0.76 (0.12)	0.21 (0.15)
East Asian faces	0.76 (0.17)	0.10 (0.07)

Western Caucasian faces	0.76 (0.17)	0.12 (0.10)

3.4.2. Eye Tracking Results

Linear Mixed Modeling analyses were conducted to test the effect of task (learning and recognition), race of face (African, East Asian, or Western Caucasian) and feature of face (eyes, mouth, or nose) on the four dependent variables: total number of first fixations, total fixation count, total fixation duration, and mean fixation duration, see Table 3.2 for a summary of statistics for each dependent variable. Analyses for total fixation duration revealed a pattern similar to those observed in the total fixation count; therefore, the results for total fixation duration were not discussed in the current chapter (refer to Appendix A for the results of total fixation duration).

Table 3.2. Summary of statistics for each dependent variable in the linear mixed models.

	Total Number of First Fixations	**Total Fixation Count**	**Total Fixation Duration**	**Mean Fixation Duration**
Task	$F(1, 5382)=.01$, $p=.912$	$F(1, 38.06)=.11$, $p=.741$	$F(1, 37.43)=5.70$, $p=.457$	$F(1, 37.44)=.001$, $p=.977$
Race	$F(2, 5382)=.01$, $p=.994$	$F(2, 5344.01)=.25$, $p=.777$	$F(2, 5343.34)=5.25$, $p=.005$*	$F(2, 61.06)=9.02$, $p<.001$**
Feature	$F(2, 5382)=439.56$, $p<.001$**	$F(2, 5344.01)=923.35$, $p<.001$**	$F(2, 5343.34)=736.71$, $p<.001$**	$F(2, 4681.75)=7.86$, $p<.001$**
Task x Race	$F(2, 5382)=.01$, $p=.986$	$F(2, 5344.01)=1.60$, $p=.202$	$F(2, 5343.34)=.05$, $p=.955$	$F(2, 1379.05)=.18$, $p=.838$
Task x Feature	$F(2, 5382)=10.88$, $p<.001$*	$F(2, 5344.01)=1.80$, $p=.165$	$F(2, 5343.34)=5.11$, $p=.006$*	$F(2, 1678.40)=2.27$, $p=.103$
Race x Feature	$F(4, 5382)=4.04$, $p=.003$*	$F(4, 5344.01)=3.51$, $p=.007$	$F(4, 5343.34)=4.80$, $p=.001$**	$F(4, 4677.44)=3.75$, $p=.005$*
Task x Race x Feature	$F(4, 5382)=1.84$, $p=.119$	$F(4, 5344.01)=1.24$, $p=.292$	$F(4, 543.34)=.48$, $p=.748$	$F(4, 4674.09)=.29$, $p=.884$

* for $p<.05$
** for $p<.001$

3.4.2.1. Total Number of First Fixations

For total number of first fixations, there was a main effect of feature, $F(2, 5382)=439.56$, $p<.001$, and interactions between task and feature, $F(2, 5382)=10.88$, $p<.001$, and race and feature, $F(4, 5382)=4.04$, $p=.003$, see Figure 3.4. Post-hoc tests using Bonferroni Pairwise Comparisons showed that participants first fixated on the nose significantly more than the eyes ($p<.001$) and mouth ($p<.001$), and the eyes significantly more than the mouth ($p<.001$).

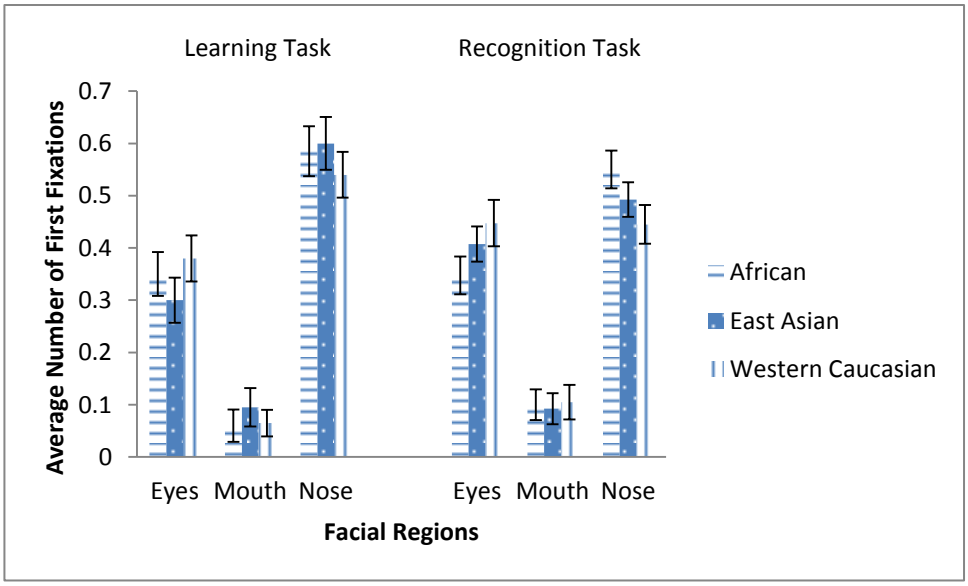


Figure 3.4. Average number of first fixations landing on the eyes, mouth, and nose during the learning and recognition tasks. Error bars report standard errors of mean. Participants tended to first fixate on the nose more than the eyes and mouth, and the eyes more than the mouth.

To examine the interaction between task and feature, separate Linear Mixed Models were conducted for each of the two tasks. Participant ID, face ID, and feature were included as factors (dependent variable = total number of first fixations; fixed factor = feature of face). There was a main effect of feature for the learning, $F(2, 1797)=210.65$, $p<.001$, and recognition, $F(2,$

3597)=266.26, p<.001, tasks. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for both tasks, participants first fixated on the eyes more than the nose (p<.001 for both tasks) and mouth (p<.001 for both tasks), and nose more than the mouth (p<.001 for both tasks). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and task were included as factors (dependent variable = total number of first fixations; fixed factor = task). There was a main effect of task for eyes, F(1, 1798)=5.62, p=.018, and nose, F(1, 1798)=10.09, p=.002, and a trend towards a main effect of task for the mouth, F(1, 1798)=3.24, p=.072. Post-hoc tests using Bonferroni Pairwise Comparisons showed that there were significantly more first fixations landing on the eyes for the recognition than learning task (p=.018) and a trend towards the mouth for the same pattern (p=.072). Conversely, there were more first fixations landing on the nose in the learning than recognition task (p=.002).

To examine the interaction between race and feature, separate Linear Mixed Models were conducted for each of three races. Participant ID, face ID, and feature were included as factors (dependent variable = total number of first fixations; fixed factor = feature of face). There was a main effect of feature for African, F(2, 1797)=184.14, p<.001, East Asian, F(2, 1797)=153.74, p<.001, and Western Caucasian faces, F(2, 1797)=135.98, p<.001. Post-hoc tests using Bonferroni Pairwise Comparisons showed that for African and East Asian faces, participants tended to direct their attention first to the nose significantly more than the eyes (p<.001 for both races of faces) and mouth (p<.001 for both races of faces), and the eyes more than the mouth (p<.001 for both races of faces). For Western Caucasian faces,

participants first fixated on the nose (p<.001) and eyes (p<.001) more than the mouth; however, the total number of first fixations on the nose and eyes did not differ significantly (p=.125). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and race were included as factors (dependent variable = total number of first fixations; fixed factor = race of face). There was a main effect of race for eyes, F(2, 1797)=3.94, p=.020, and nose, F(2, 1797)=4.43, p=.012, but not for the mouth, F(2, 1797)=.088, p=.916. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for the eyes, there were significantly more first fixations for Western Caucasian than African faces (p=.019). However, the number of first fixations on the eyes for East Asian faces did not differ significantly from Western Caucasian (p=.171) and African faces (p=1). As for the nose, there were significantly more first fixations for African than Western Caucasian faces (p=.01). However, the number of first fixations on the nose for East Asian faces did not differ significantly from Western Caucasian (p=.219) and African faces (p=.741).

3.4.2.2. Total Fixation Count

For total fixation count, there was a main effect of feature, F(2, 5344.01)=923.35, p<.001 and an interaction between race and feature, F(4, 5344.01)=3.51, p=.007, see Figure 3.5. Post-hoc tests using Bonferroni Pairwise Comparisons showed that participants fixated on the eyes significantly more than the nose (p<.001) and mouth (p<.001), and the nose more than the mouth (p<.001), see Figure 3.6 for a heat map of participants' fixation distribution.

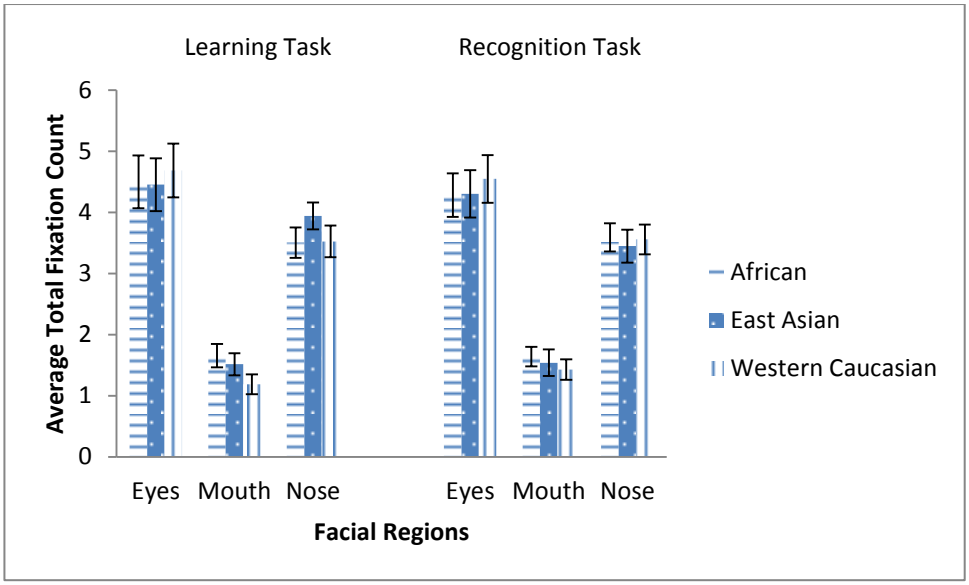


Figure 3.5. Average total fixation count for the eyes, mouth, and nose during the learning and recognition tasks. Error bars report standard errors of mean. Participants fixated more on the eyes, followed by the nose, then mouth.

To examine the interaction between race and feature, separate Linear Mixed Models were conducted for each of the three races. Participant ID, face ID, and feature were included as factors (dependent variable = total fixation count; fixed factor = feature of face). There was a main effect of feature for African, $F(2, 1797)=269.16$, $p<.001$, East Asian, $F(2, 1797)=285.11$, $p<.001$, and Western Caucasian faces, $F(2, 1797)=371.71$, $p<.001$. Post-hoc tests using Bonferroni Pairwise Comparison revealed that for all three races of faces, participants made significantly more fixations on the eyes than the nose ($p<.001$ for all three races) and mouth ($p<.001$ for all three races), and the nose more than the mouth ($p<.001$ for all three races). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and race were included as factors (dependent variable = total fixation count; fixed factor = race of face). There was a main effect of race for the mouth, $F(2, 1797)=7.12$, $p=.001$, but not the eyes, $F(2, 1797)=1.59$, $p=.205$,

and nose, $F(2, 1797)=.16$, $p=.849$. Post-hoc tests using Bonferroni Pairwise Comparisons showed that for the mouth, there were significantly more fixations landing on African than Western Caucasian faces ($p=.001$) and a trend towards more fixations landing on East Asian than Western Caucasian faces ($p=.062$). However, the number of fixations landing on African and East Asian faces did not differ significantly ($p=.467$)

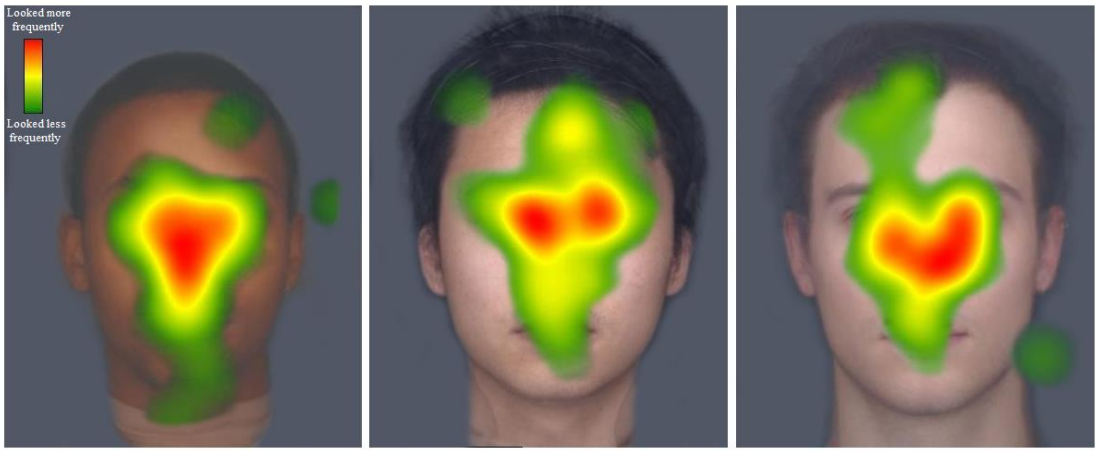


Figure 3.6. The colored areas show locations where Malaysian Chinese participants focused on when recognizing faces. Composite images are shown for illustration purposes. Real faces were used in the actual experiment.

3.4.2.3. Mean Fixation Duration

For mean fixation duration, there was a main effect of race, $F(2, 61.06)=9.02$, $p<.001$, and feature, $F(2, 4681.75)=7.86$, $p<.001$, and an interaction between race and feature, $F(4, 4677.44)=3.75$, $p=.005$, see Figure 3.7. For race, post-hoc tests using Bonferroni Pairwise Comparisons showed that participants made fixations of significantly longer duration for African than East Asian ($p=.001$) and Western Caucasian faces ($p=.002$). However, the duration of fixations for East Asian and Western Caucasian faces did not differ significantly ($p=1$). As for feature, post-hoc tests using Bonferroni

Pairwise Comparisons showed that participants' duration of fixations was longer for the mouth (p=.002) and eyes (p=.002) than the nose. However, the duration of fixations for the mouth and eyes did not differ significantly (p=1).

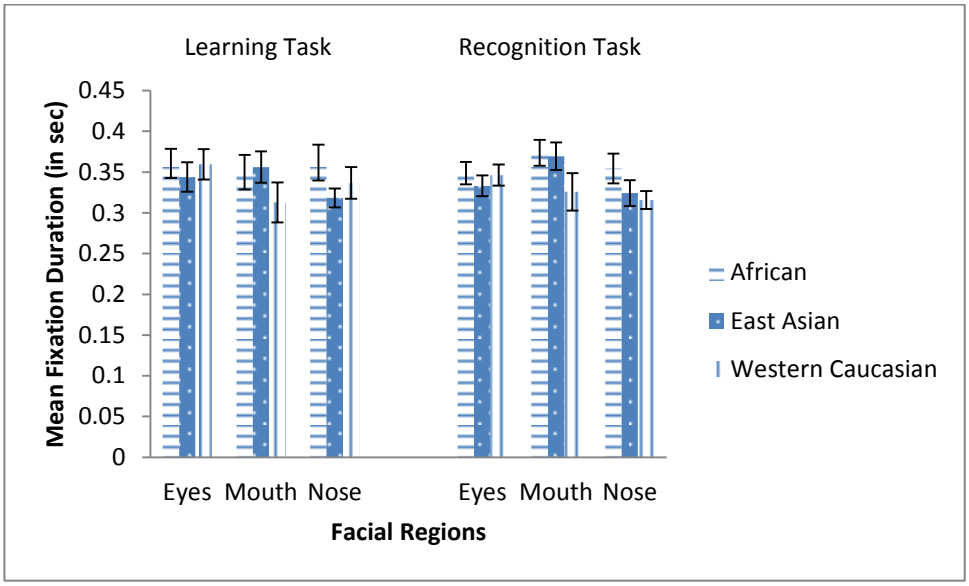


Figure 3.7. Mean fixation duration on the eyes, mouth, and nose during the learning and recognition tasks, in seconds. Error bars report standard errors of mean. The mean fixation duration was longer for the mouth and eyes than the nose.

To examine the interaction between race and feature, separate Linear Mixed Models were conducted for each of the three races. Participant ID, face ID, and feature were included as factors (dependent variable = mean fixation duration; fixed factor = feature of face). There was a main effect of feature for East Asian, $F(2, 1588)=10.42$, $p<.001$ and Western Caucasian faces, $F(2, 1564)=6.49$, $p=.002$, but not for African faces, $F(2, 1612)=.26$, $p=.775$. Post-hoc tests using Bonferroni Pairwise Comparisons showed that for East Asian faces, the mean fixation duration for the mouth was significantly longer than the eyes ($p=.048$) and nose ($p<.001$), and there was a trend towards significantly longer mean fixation duration for the eyes than the nose ($p=.067$).

For Western Caucasian faces, the mean fixation duration for the eyes was significantly longer than the nose (p=.001); however the mean fixation duration for the mouth was not significantly shorter than the eyes (p=.486) and not significantly longer than the nose (p=.175). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and race were included as factors (dependent variable = total number of first fixations; fixed factor = race of face). There was a main effect of race for eyes, F(2, 1705)=3.06, p=.047, mouth, F(2, 1334)=3.16, p=.043, and nose, F(2, 1725)=12.71, p<.001. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for the eyes, there was a trend towards a longer duration of fixations for African than East Asian faces (p=.05). However, the mean fixation duration on the eyes for Western Caucasian faces did not differ from African (p=1) and East Asian faces (p=.257). For the mouth, the duration of fixations was significantly longer for African than Western Caucasian faces (p=.047); however, the mean fixation duration on the mouth for East Asian faces was not significantly different from African (p=1) and Western Caucasian faces (p=.192). As for the nose, the mean fixation duration was significantly longer for African than Western Caucasian (p<.001) and East Asian faces (p<.001). However, the mean fixation duration for the mouth did not differ significantly for Western Caucasian and East Asian faces (p=1).

3.5. Discussion

The current study examined Malaysian Chinese participants' recognition sensitivity and eye movement strategies when viewing East Asian, Western Caucasian, and African faces. The primary findings were that 1) Malaysian Chinese participants

performed equally well at recognizing East Asian and Western Caucasian faces, but less well at recognizing African faces, 2) when perceiving all three races of faces, participants tended to fixate first on the nose more than the eyes and mouth, 3) Malaysian Chinese participants displayed a combination of Eastern and Western eye movement strategies with most fixations clustered around the eyes, followed by the nose then mouth, and finally 4) the mean fixation duration for the eyes and mouth was significantly longer than the nose.

Recent face recognition studies involving eye tracking discovered that East Asian observers fixated more on the nose, an optimal strategy to integrate facial feature information holistically, while Western Caucasians perceive faces analytically by moving between the eyes and mouth (Blais et al., 2008). However, the current study involving Malaysian Chinese participants reported a combination of Eastern and Western eye movement strategies employed during face perception, with more fixations landing on the eyes than the nose, followed by the mouth. To ensure that the current eye tracking result is comparable with a previous face recognition study (Blais et al., 2008), the AOI sizes were reduced to fit tightly around the eyes, nose, and mouth. Eye tracking result for the new set of AOIs revealed that Malaysian Chinese participants consistently employed the same fixation pattern (with more fixations on the eyes, followed by the nose, then mouth) regardless of the size of AOIs.

Malaysian Chinese first fixated on the nose during the learning task, indicating that a majority of the participants instinctively use a holistic approach to make first impressions or subjective judgments (i.e. the attractiveness) about a face. However, there were more first fixations on the eyes during the recognition task, suggesting that participants are more likely to rely on the individuating feature of the face when recalling identities. Analyses on the total number of first fixations also showed a

preference towards Western Caucasian over African eyes and African over Western Caucasian nose. This implies that when viewing faces of different races, at first glance, participants may rely on different features to make judgments about them.

In the learning and recognition tasks, the total fixation count was higher for the eyes, followed by the nose then mouth, implying that Malaysian Chinese tended to adopt a mixture of Eastern and Western fixation patterns when perceiving faces. Participants fixated more on the mouth for African and East Asian than Western Caucasian faces. Attention to the mouth for African faces conjuncts with previous findings that different facial features are relevant to different races of faces; the lower facial features were found to be the diagnostic features for African faces (Ellis et al., 1975; Hills & Lewis, 2006).

However, the mean fixation duration revealed that the eyes and mouth rendered longer fixations than the nose, implying that participants had fewer but longer fixations on the mouth and more short fixations on the nose. The interaction between race and feature showed that for East Asian faces, the mean fixation duration on the mouth was significantly longer than the eyes than the nose, whereas for Western Caucasian faces, participants made fixations of significantly longer duration on the eyes than the nose and a trend towards longer duration for the mouth than the nose; however, the mean fixation duration for the eyes and mouth was not significantly different. Participants may have been particularly drawn to the eyes for Western Caucasian faces possibly because they convey more individuating information (e.g. eye color) of a face than for East Asian and African faces, which tend to have rather uniform eye color. Additionally, the mean fixation duration showed that participants tended to fixate for longer durations on African than East Asian and Western Caucasian faces, especially on the mouth and nose. The longer

durations for African faces may be due to the lack of familiarity, suggesting that the processing of African faces is more effortful for Malaysian Chinese participants. The lower facial features of African faces not only rendered more fixations but also longer mean fixation duration, providing stronger evidence that the lower facial features may be especially informative for identifying and making subjective judgments of African faces.

Perhaps the diversity of the Malaysian culture plays a large role in shaping the unique fixation patterns that Malaysian Chinese participants adopt. As compared with the Chinese or Japanese population, who are largely influenced by their respective culture, Malaysian Chinese have greater exposure to the West as reflected in the high percentage of English radio stations and Western movies shown in Malaysian cinema (Davies, 2011; Epstein, 2011). The population of Malaysia is also highly diverse, with Malays composing 50.4% of the population, Chinese 23.7%, indigenous 11%, Indian 7.1%, and others composing 7.8% (including Western Caucasians and Africans; "The World Factbook," 2011). In addition, Malaysian Chinese would be more familiar with Western Caucasian than African faces as the Department of Immigration in 2006 listed the United Kingdom as the fourth largest source of expatriates to Malaysia while no African (or black majority) country was listed as a major source of immigrants in any class of international migration (Kanapathy, 2008). All of the participants also attend a branch campus of a British university, which increases their exposure to Western people and culture. Familiarity with the Western culture may have influenced the fixation pattern that Malaysian Chinese employ to best recognize the faces which they most frequently encounter. Rather than adopting the Eastern or Western fixation pattern, Malaysian Chinese participants employed a mixed strategy that is strikingly different from either of these fixation patterns previously reported

(Blais et al., 2008), but perhaps more aligned with those of British born Chinese (Kelly et al., 2011a). Of particular note in this study, was the low number of fixations directed at the mouth, which rendered the eye movement patterns of Malaysian Chinese observers distinct from both the East Asian and West Caucasians reported in previous studies (Blais et al., 2008).

The combination of Eastern and Western strategies proved advantageous for Malaysian Chinese participants to accurately recognize East Asian and Western Caucasian faces. Although the less familiar African faces were not recognized as well, participants were more attentive to the diagnostic features of African faces (i.e. the lower facial features). Previous studies have shown that people of different races use different facial features to describe (Ellis et al., 1975) and recognize (Blais et al., 2008) faces. Furthermore, research has suggested that the other-race effect may be affected by the looking strategy employed when viewing other-race faces. Training observers to focus on the critical features relevant to a specific race of face (e.g. eyes for Caucasians and lips for Africans) has been found to reduce the other-race effect, implying that observers may have been using the wrong looking strategy to perceive other-race faces, which resulted in poorer recognition performance (Hills & Lewis, 2006). In line with this finding, attention to the eyes appeared to be associated with Malaysian Chinese participants' ability to recognize Western Caucasian faces.

The results of this chapter suggest that Malaysian Chinese participants learned to use a fixation pattern that is advantageous for them to recognize the multicultural faces that they encounter; however, it is unknown whether Malaysian Chinese observers would consistently employ the same hybrid of Eastern and Western strategies across various tasks and whether this strategy would enable observers to perform equally well on other tasks (i.e. categorizing facial expressions of emotion).

The next chapter aimed to identify the eye movement strategy that Malaysian Chinese participants employ when viewing photos of East Asian and Western Caucasian models posing Ekman's six universal emotions ("happy," "surprise," "fear," "disgust," "anger," and "sad"; 1970) plus "neutral," and participants' performance on an emotion recognition task.

Chapter Four

Static Emotion Recognition

Abstract

Facial expressions of emotion have long been considered an innate, evolved behavior that is universal. However, cross-cultural studies have found systematic differences in East Asian and Western Caucasian observers' perception of facial expressions, challenging the universality hypothesis. Recent eye tracking studies, for instance, found that East Asians recognize facial expressions by focusing on the eye region alone whereas Western Caucasians look evenly across the face by moving between the eyes and mouth. East Asian observers' lack of focus on the mouth may have contributed to the misrecognition of certain expressions, such as "fear" and "disgust." The current study aimed to examine whether Malaysian Chinese participants continue to employ an intermediate strategy or shift towards an Eastern or Western pattern (by fixating more on the eyes or mouth respectively) during an emotion recognition task. Results showed that participants fixated most on the eyes, followed by the nose then mouth when recognizing facial expressions of emotion. Possibly due to the lack of focus on the mouth, "fear" was recognized significantly worse than other emotions. As compared to the face recognition task, fixations on the eye area were reduced and fixations on the nose and mouth areas were increased for the emotion recognition task. The current results suggest that the looking pattern an individual adopts may be influenced by the culture and environment that one is familiar with.

4.1. Introduction

The mutual understanding of emotion through the exchange of facial expressions is the foundation of human social interaction. Facial expressions not only bear information about an individual's emotional state but also convey non-verbal cues about one's motivations, social intentions, and interpersonal traits (Darwin, 1872; Knutson, 1996). Therefore, the ability to recognize and respond to others' display of emotions is vital in our daily social encounters. Although facial expressions of emotion are long thought to be universal, cross-cultural studies have found conflicting evidence that challenges the universality hypothesis (Birdwhistell, 1970; Klineberg, 1940; Biehl, Matsumoto, Ekman, Hearn, Heider, Kudoh, & Ton, 1997; Moriguchi, Ohnishi, Kawachi, Mori, Hirakata, Yamada, Matsuda, & Komaki, 2005). Researchers have attempted to reconcile the two extreme ends of the argument by differentiating elements of facial expressions that are universal and culture specific. In the following sections, the theoretical framework of facial expressions of emotion and the identification of important facial cues used to categorize emotions will be discussed in more detail.

4.1.1. Theoretical Framework

Early studies on facial expressions of emotion suggested that all humans display the six basic emotional states ("happy," "surprise," "fear," "disgust," "anger," and "sad") using the same facial movements that are biologically and evolutionarily determined, supporting the universality hypothesis (Darwin, 1872). Researchers argued that the meanings of facial expressions would have diverged if they were merely cultural traits that are socially learned. Numerous studies involving diverse populations have provided evidence supporting the universality hypothesis thereafter.

For instance, studies found that New Guineans, although minimally exposed to the mass media presentation of literate cultures, were able to categorize most facial expressions accurately except for "fear" (Ekman & Friesen, 1971). The authors suggested that New Guineans consistently confused "fear" with "surprise" possibly because of the cultural interpretation of fearful situations as surprising. A recent study examining spontaneous display of emotions in blind individuals also found no difference between congenitally blind, noncongenitally blind, and sighted athletes, which suggests that the display of emotions is not learned and modeled from observation but are genetically coded for all humans (Matsumoto & Willingham, 2009).

Conversely, the culture specificity hypothesis suggests that facial expressions of emotion are different in each culture. Researchers proposed that the display of emotion, which act as a means of communication, are composed of units and organized similar to the components of spoken language (Birdwhistell, 1970). Although certain patterns of behavior, such as laughing and crying, are universal, the expressions of emotion like anger and fear are not. Hence, facial expressions and body movements would be socially learned according to the structure of individual society (Klineberg, 1940). Studies using subjective ratings found that while observers from 7 different countries agreed on the facial expressions for each photograph, the level of agreement between cultures differed, implying that the semantic or affective meanings of each emotion may vary across cultures (Biehl et al., 1997). Further examination of observers' intensity ratings suggested dissimilarity in either the portrayal or judgment of facial expressions. Brain imaging study also showed that when viewing photographs of "happy" and "fearful" own- and other-race faces, Caucasian and Japanese participants responded differently to "fearful" faces,

suggesting that the perception of facial expressions of emotion may be influenced by individual culture (Moriguchi et al., 2005).

To reconcile the two extreme ends of the argument, Ekman (1970) differentiated elements of facial expressions that are universal and culture specific, proposing that the expressions of emotion might vary across cultures as a result of three main factors. Firstly, certain stimuli are learned to be the elicitors of specific emotions in different cultures (e.g. a husband's second marriage may evoke a wife's jealousy in one culture but may be a stimulus for pride in another); therefore, the comparison of a same event in two cultures may not be accurate, as it does not necessarily elicit an identical emotion. Secondly, different cultures may have different display rules that govern one's display of emotions in particular social settings (e.g. even if the same event evokes joyous emotion in two cultures, one culture may de-intensify joy with a modest smile while another may over-intensify the emotion by jumping and shouting in joy). Lastly, the behavioral consequences of facial expressions of emotion may vary across cultures (e.g. an individual may choose to walk away from a situation to cope with overwhelming feelings of anger, fear, or even disgust); thus, the body postures and movements following a facial expression does not necessarily reflect the actual emotion felt. Ekman (1970) argued that although the evoking stimuli, display rules, and behavioral consequence of emotion may differ across cultures, the facial expressions of the primary emotions are universal to mankind.

4.1.2. Identification of Important Facial Cues Used to Categorize Emotions

Recent studies have aimed to identify the facial cues that participants utilize to categorize facial expressions of emotion. In a cross-cultural study utilizing eye

tracking methodology, East Asian and Western Caucasian participants were required to categorize own- and other-race faces displaying Ekman's six universal emotions ("happy," "surprise," "fear," "disgust," "anger," and "sad"; 1970) plus "neutral" (Jack, Blais, Scheepers, Schyns, & Caldara, 2009). Behavioral results showed that East Asian observers mistakenly identified "fear" and "disgust" as "surprise" and "anger;" conversely, Western Caucasian observers recognized all facial expressions with relatively high accuracy. When categorizing facial expressions of emotion, East Asians focused on the eye region alone whereas Western Caucasians look evenly across the face by moving between the eyes and mouth. Perhaps due to the lack of focus on the mouth, East Asian observers may have misidentified relevant information, which led to the miscategorization of certain emotions, such as "fear" and "disgust." When in doubt of their answers, East Asian observers tended to choose a less threatening emotion, such as "surprise." These findings indicate that certain emotions may be expressed differently in the East as observers from this culture have learned to adopt a different fixation pattern to categorize facial expressions.

A more recent eye tracking study investigated the effect of expression intensity on female participants' categorization accuracy and eye movement strategy (Guo, 2012). Twenty-eight Western Caucasian faces posing for "neutral" and six high intensity facial expressions of "happy," "surprise," "fear," "disgust," "anger," and "sad" were collected to produce 5 levels of intensity, which ranged from 20% to 100% with 20% increments, by morphing the neutral face with the 6 expressions. Participants, who were also Western Caucasians, were required to categorize the 120 images with varying intensities of expressions by pressing a button (for the collection of reaction time data) before giving a verbal response of their answers. Results showed that although participants fixated least (in terms of fixation number and

fixation duration) on “happy” faces, the categorization accuracy was the highest for the emotion; on the other hand, participants fixated most on “fearful” faces but showed the lowest categorization accuracy. These differences observed in “happy” and “fearful” faces were present at the lowest expression intensity of 20% and were consistent throughout the range of intensities. In expressions of lower intensities, participants tended to confuse “surprise,” “fear,” “anger,” and “disgust” with emotions like “sad” and “happy,” possibly because “sad” and “happy” are relatively common expressions encountered in everyday interactions. For expressions of medium and high intensity, participants consistently misidentified “fear” as “surprise and “disgust” as “anger.” The author proposed that the confusion between facial expressions may be due to the shared muscle action units between the confused expressions, the lack of exposure to these expressions, and the differences in saliency of the expressions. Participants were better and quicker at recognizing facial expressions of emotion with higher intensity; moreover, the number of fixations was reduced when viewing expressions of higher intensity. The effect was most evident in intensity up to 60%; however, intensity above 60% did not have further impact on participants’ reaction time and fixations. Further analyses of the eye tracking data revealed that participants consistently fixated more and longer on the eyes, followed by the nose then mouth across all expressions and intensities. Although participants generally utilized a holistic approach to view all expressions, participants tended to direct more attention to certain regions of the face for different expressions; more specifically, when comparing individual feature across expressions, participants fixated more on the mouth for “happy” faces, on the nose for “disgust” and “sad” faces, and on the eyes for “angry,” “fear,” and “surprise” faces. The fixation pattern identified in Guo’s (2012) study is different from those observed in a previous eye

tracking study (Western Caucasian participants made even distribution of fixations on the eyes and mouth; Jack et al., 2009), possibly because expressions with lower intensities may be more ambiguous; hence, gaining a more holistic view of the face would allow better categorization of the expressions.

Further research aimed to identify the different facial cues that Japanese and American observers rely on to categorize facial expressions of "happy" and "sad" in illustrated faces and edited facial expressions from real people (Yuki, Maddux, & Masuda, 2007). In both sets of images, cues from the eyes and mouth were manipulated to produce different combinations of "happy" and "sad" faces (e.g. "happy" eye and "sad" mouth or "sad" eye and "happy" mouth). Results revealed that people from different cultures rely on different facial cues to interpret facial expressions of emotion; Japanese observers weighted cues displayed in the eyes more than Americans whereas Americans weight cues in the mouth more when categorizing expressions. The authors proposed that people learn to rely on different facial cues because of the cultural differences in how the emotions are expressed. In the West, where individualism is embraced, people are encouraged to express their true feelings as it reflects the acceptance of one's true self (Heine, Lehman, Markus, & Kitayama, 1999). In contrast, Eastern cultures value the control of emotion to maintain harmonious relationships with others (Friesen, 1972). These cultural differences in one's social environment may affect how emotions are expressed. Furthermore, it has been shown that the muscles around the eyes are more difficult to control than the muscles around the mouth, suggesting that the eyes may be a more useful cue to identifying one's true emotions (Ekman, 1992). Hence, in a culture where emotions may be suppressed, Japanese observers learn to focus on the eye region as it provides more diagnostic information of another's true emotions, while in

America where the expression of true emotions is encouraged, people focus on the most expressive region of the face—the mouth.

A study employing the Bubbles technique also found that Caucasian participants tended to utilize cues from the mouth area to categorize facial expressions of emotion (Blais, Roy, Fiset, Arguin, & Gosselin, 2012). Visual information within a stimulus is randomly sampled and each subset of this information is presented to the participant on each individual trial, allowing researchers to identify which subset of information was used to achieve categorization of each facial expression, see Figure 4.1 for an example of how the bubblized stimulus was created. Participants completed 4000 trials, in which bubblized images or videos of eight facial expressions (i.e. "neutral," "happy," "surprise," "fear," "disgust," "anger," "sad," and "pain") were presented for 500 milliseconds each, followed by participants' response with a corresponding key press. Results showed that participants used facial cues from the mouth for both static and dynamic stimuli. Although previous eye tracking studies have found that Western Caucasian participants distributed even fixations on the eyes and mouth (Jack et al., 2009), the authors suggested that the number of fixations may not necessarily reflect the actual visual information processed. More specifically, the eyes are represented by higher spatial frequency, which allow participants to process both the eye and mouth region when fixating on the eyes; however, fixations on the mouth allow processing of the mouth area only. Therefore, even though the fixations landing on the eyes and mouth were similar, the mouth may, in fact, reveal more useful information than the eyes.

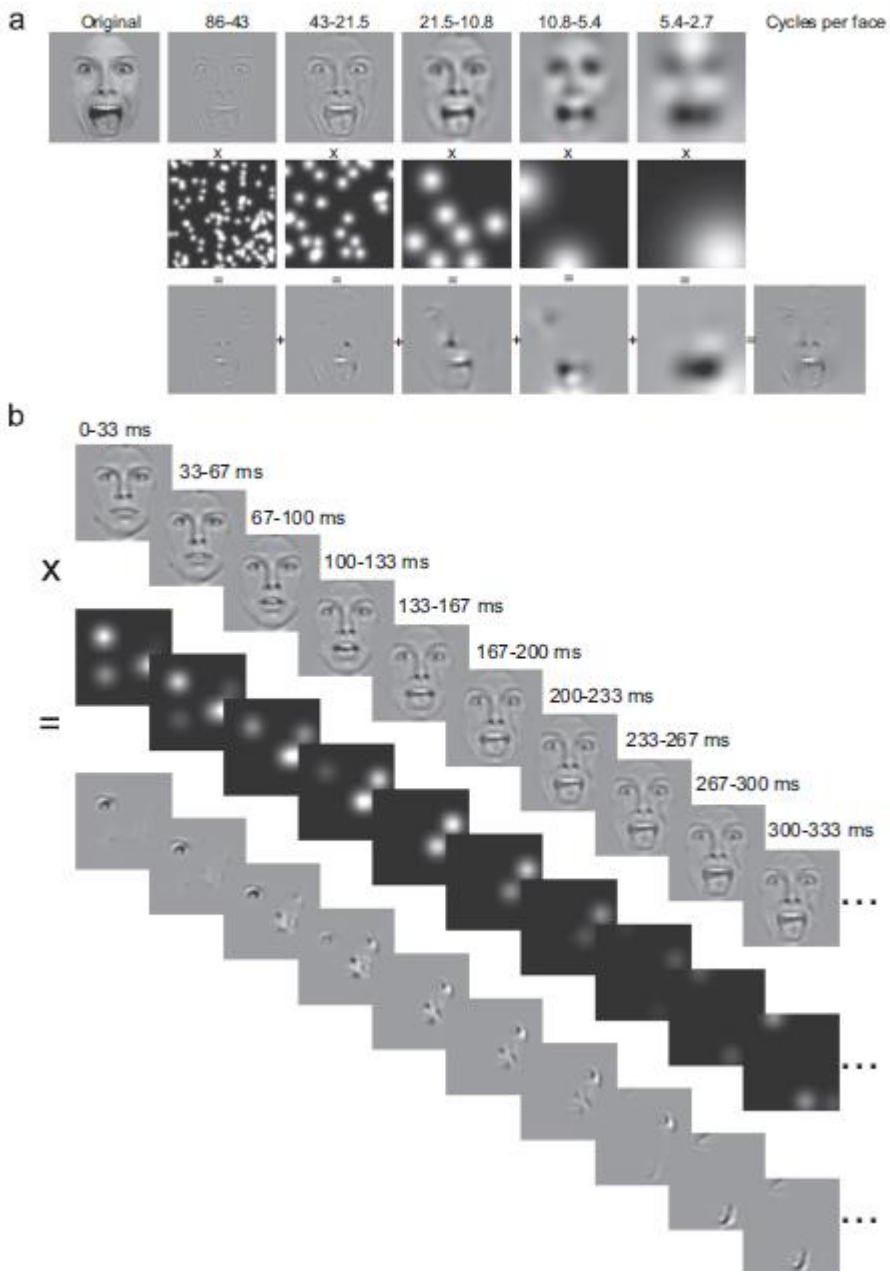


Figure 4.1. Examples of the creation of bubblized image (a) and video (b) used Blais et al. (2012)'s study.

4.2. Current Study

The research reviewed so far suggests that there are distinct Eastern and Western viewing patterns when identifying facial expressions of emotion (Jack et al., 2009). However, as presented in Chapter Three, face recognition studies have

revealed that exposure to and familiarity with another culture may foster a hybrid fixation pattern that enables observers to recognize both own- and other-race faces (Kelly, Jack, Miellet, De Luca, Foreman, & Caldara, 2011; Tan, Stephen, Whitehead, and Sheppard, 2012). This raises the possibility that hybrid looking strategies may also be used for identifying facial expressions in such societies.

The current study aimed to examine Malaysian Chinese observers' sensitivity and eye movement strategy in categorizing East Asian and Western Caucasian facial expressions. If Malaysian Chinese follow an Eastern looking pattern, participants would focus mainly on the eye region when recognizing facial expressions. This might also result in better identification for own- than for other-race faces as eyes may be a more accurate cue for recognizing emotions in East Asian faces (Yuki et al., 2007). However, because of the lack of attention to the mouth, Malaysian Chinese might be poor at recognizing certain emotions, like "fear" or "disgust" (Ekman & Friesman, 1971). On the other hand, if Malaysian Chinese are affected by Western culture, we might see a hybrid of Eastern and Western strategies; with fixations landing on the eyes and mouth more than the nose. A combination of both strategies may enable participants to identify expressions from same- and other-race faces with equal success.

4.3. Methods

4.3.1. Participants

To collect photographs for the emotion recognition study, "models" (term used as short-hand—the volunteers were not trained models) were recruited to pose for Ekman's 6 universal emotions ("happy," "surprise," "fear," "disgust," "anger,"

and “sad”; 1970) plus “neutral.” Models were 11 male, 8 female East Asian (mean age 20.75 years) and 7 male, 10 female Western Caucasian (mean age 30.75 years) students or staff in the University of Nottingham Malaysia Campus who participated voluntarily.

In the validation phase, 44 University of Nottingham Malaysia Campus students (9 males, 17 females, mean age 22.67 years) were asked to categorize the facial expressions to ensure that the photographs portrayed the emotions assigned. Participants were of different racial backgrounds (e.g. African, East Asian, Western Caucasian, etc.) who participated voluntarily.

Twenty-six Malaysian Chinese students attending the University of Nottingham Malaysia Campus (20 males, 24 females, mean age 21.35 years) participated in the emotion recognition study. All participants have not lived outside of Malaysia for more than three years. All participants had normal or corrected vision and were given a bar of chocolate for their participation.

Written informed consent was obtained from participants in all three phases and the protocol was approved by the University of Nottingham, Faculty of Science Ethics Committee.

4.3.2. Materials

Prior to the main experiment, 36 sets of images (11 male, 8 female East Asian, mean age 20.75 years; 7 male, 10 female Western Caucasian, mean age 30.75 years) displaying Ekman’s 6 universal emotions (“happy,” “surprise,” “fear,” “disgust,” “anger,” and “sad”; 1970) plus “neutral” were collected. The images were captured using the Canon EOS 550D with an aperture of 5.0f, shutter speed of 1/100second,

and an ISO of 400. Multiple photographs were taken of each emotion for each model and a total of 369 images, displaying one of seven emotions, were used in the validation phase. All images were photographed in a Munsell N5 neutral gray painted lighting booth illuminated with d65 fluorescent tubes, in high-frequency fixtures to reduce the effects of flicker, with no other lightings present in the room (Verivide, UK). Images were aligned to the eyes and cropped around the face using Psychomorph software (Tiddeman & Perrett, 2001). The images were 270x333 pixels in size and were presented at a distance of about 60cm on a HP desktop PC using E-Prime. Forty-four University of Nottingham Malaysia Campus students (9 males, 17 females, mean age 22.67 years) of different racial backgrounds (e.g. African, East Asian, Western Caucasian, etc.) were asked to categorize the facial expressions by pressing a corresponding response key on the keyboard.

Participants' responses were compiled and the percentage of accuracy for each of the 369 images was calculated. Out of the 36 models, only 8 models whose facial expressions were identified with an average accuracy of 70 percent or higher were selected for the main experiment; totaling 56 images (i.e. 2 male, 2 female East Asians and 2 male, 2 female Western Caucasians, each displaying 7 emotions; see Figure 4.2). The images were presented at a distance of approximately 60cm on a Tobii T60 eye tracker (a 17in TFT monitor with a screen resolution of 1280x1024 pixels), using Tobii Studio software, version 2.3.



Figure 4.2. Examples of stimuli used in the experiment. Model displaying the six universal emotions: “happy,” “surprise,” “fear,” “disgust,” “anger,” and “sad.”

4.3.3. Procedure

The main experiment was carried out on a Tobii eye tracker. On each of the 56 trials, a central fixation cross was presented for one second followed by a face presented pseudorandomly in one of four quadrants of the computer screen to avoid fixation biases. The face stimulus displaying one of the seven expressions was presented for 5 seconds, after which participants were required to identify the most appropriate emotion via a seven-alternative forced-choice procedure by pressing a corresponding response key on the keyboard. Each response was followed by the central fixation cross, which preceded the next face stimulus. At the end of the

experiment, participants were asked to fill out the Autism Quotient questionnaire, which is a 50 item forced-choice self-report questionnaire that assesses autistic traits.

4.3.4. Eye Tracking

Eye tracking information and the calibration procedure were identical to those presented in Chapter Three.

4.3.5. Data Analyses

To determine participants' categorization sensitivity, a-prime (A') values were used to account for any bias of participants to pick a certain facial expression label over others by taking into account both hits (i.e. correct detection of an expression as displayed by the image [e.g. identifying happy as happy]) and false alarms (i.e. incorrect identification of a different expression [e.g. surprise] as the expression in question [e.g. happy]). Refer to Chapter Three for a detailed description of A'.

Tobii Studio software was utilized to process raw data directly from the eye-tracker. The average number of fixations landing on three predefined areas (i.e. eyes, mouth, and nose) was calculated using an area of interest (AOI) analysis, see Figure 4.3. Fixations were defined as two or more consecutive samples falling within a 35 pixel radius and data was sampled approximately every 17 milliseconds. Each participant had to complete 56 trials for 5 seconds each. To ensure the validity of the eye tracking data, only participants who had an average number of fixations equal to or more than 5 were included in the data set. One female participant did not meet the criteria and was excluded from the eye tracking analyses.

Statistical analyses were performed using SPSS version 21. Linear Mixed Modeling analyses were conducted to test the effect of facial expressions of emotion

(“neutral,” “happy,” “surprise,” “fear,” “disgust,” “anger,” and “sad”), race of face (East Asian or Western Caucasian) and feature of face (eyes, mouth, or nose) on the four dependent variables: total number of first fixations (i.e. the accumulated number of times that the first fixation made on a face fell on either the eyes, mouth, or nose for all trials), total fixation count (i.e. the accumulated number of fixations landing on the eyes, mouth, and nose for all trials), total fixation duration (i.e. the accumulated fixation duration (in seconds) landing on the eyes, mouth, and nose for all trials), and mean fixation duration (i.e. total fixation duration divided by total fixation count). Participant ID, face ID, emotion, race of face, feature of face were included as factors. All main effects and the two- and three-way interactions between emotion, race of face, and feature of face were included in the model (dependent variable = total number of first fixations, total fixation count, total fixation duration and mean fixation duration; fixed factors = emotion, race of face, and feature of face). Participant ID and face ID were included as subjects to avoid pseudoreplication. Face ID nested within race was included as a random factor for total fixation count and mean fixation duration. No random factors were included for total number of first fixations and total fixation duration because the final Hessian matrix was not positive, which casts doubt on the validity of results; however, all results with or without face ID nested within race included as a random factor were identical.

To compare how participants differed in their looking patterns when performing different tasks, eye tracking data for the current experiment was also analyzed alongside the data from the face recognition experiment presented in Chapter Three, which explored the eye movement strategies that participants used when recognizing identities.

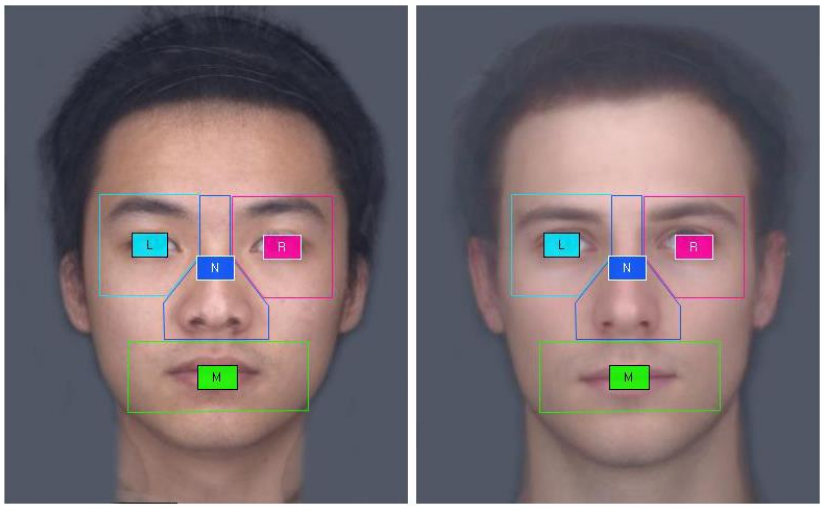


Figure 4.3. The predefined area of interest (AOI) used to analyze eye gaze. Composite images are shown for illustration purposes. Real faces were used in the actual experiment.

4.4. Results

4.4.1. Categorization Sensitivity

One-sample t-tests on A' values showed that Malaysian Chinese participants were able to recognize East Asian ($t(25)=37.52$, $p<.001$ for "neutral," $t(25)=453.53$, $p<.001$ for "happy," $t(25)=76.22$, $p<.001$ for "surprise," $t(25)=13.64$, $p<.001$ for "fear," $t(25)=35.63$, $p<.001$ for "disgust," $t(25)=34.14$, $p<.001$ for "anger," and $t(25)=39.94$, $p<.001$ for "sad") and Western Caucasian facial expressions ($t(25)=51.17$, $p<.001$ for "neutral," $t(25)=157.77$, $p<.001$ for "happy," $t(25)=16.08$, $p<.001$ for "surprise," $t(25)=15.69$, $p<.001$ for "fear," $t(25)=35.14$, $p<.001$ for "disgust," $t(25)=28.35$, $p<.001$ for "anger," and $t(25)=40.65$, $p<.001$ for "sad") above chance level (test value = 0.5).

Table 4.1. Hit (H) rates and false alarm (FA) rates in recognizing East Asian (EA) and Western Caucasian (WC) facial expressions of "neutral," "happy," "surprise," "fear," "disgust," "anger," and "sad" faces.

	H Rates (SD) for EA faces	FA Rates (SD) for EA faces	H Rates (SD) for WC faces	FA Rates (SD) for WC faces
Neutral	0.72 (0.18)	0.04 (0.03)	0.75 (0.15)	0.03 (0.02)
Happy	0.90 (3.4E-16)	0.03 (0.02)	0.89 (0.04)	0.05 (0.03)
Surprise	0.86 (0.08)	0.09 (0.04)	0.66 (0.20)	0.13 (0.05)
Fear	0.36 (0.21)	0.06 (0.05)	0.35 (0.20)	0.07 (0.06)
Disgust	0.79 (0.14)	0.11 (0.05)	0.80 (0.14)	0.10 (0.06)
Anger	0.67 (0.21)	0.02 (0.01)	0.62 (0.23)	0.03 (0.01)
Sad	0.72 (0.17)	0.03 (0.02)	0.74 (0.20)	0.02 (0.01)

A 2 (Race of Face: East Asian or Western Caucasian) x 7 (Emotions: "neutral," "happy," "surprise," "fear," "disgust," "anger," and "sad") one-way repeated measures analysis of variance (ANOVA) was conducted on participants' sensitivity (A') in recognizing facial expressions of emotion, see Table 4.1 for participants' hit rates and false alarm rates. There was a main effect of both race of face, $F(1, 25)=4.95$, $p=.035$ and facial expressions of emotion, $F(3.5, 87.41)=47.18$, $p<.001$ were found, both Greenhouse-Geisser corrected. Participants categorized East Asian expressions significantly better than Western Caucasian ones. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that participants were best at recognizing "happy" ($p=.001$ for "sad" and $p<.001$ for all 5 other comparisons) while "fear" was categorized significantly worse than all other expressions ($p=.001$ for all 6 comparisons). "Fear" was not only the least recognized emotion, but also the only one consistently confused with another emotion; in 50% of the trials, participants mistook "fear" for "surprise."

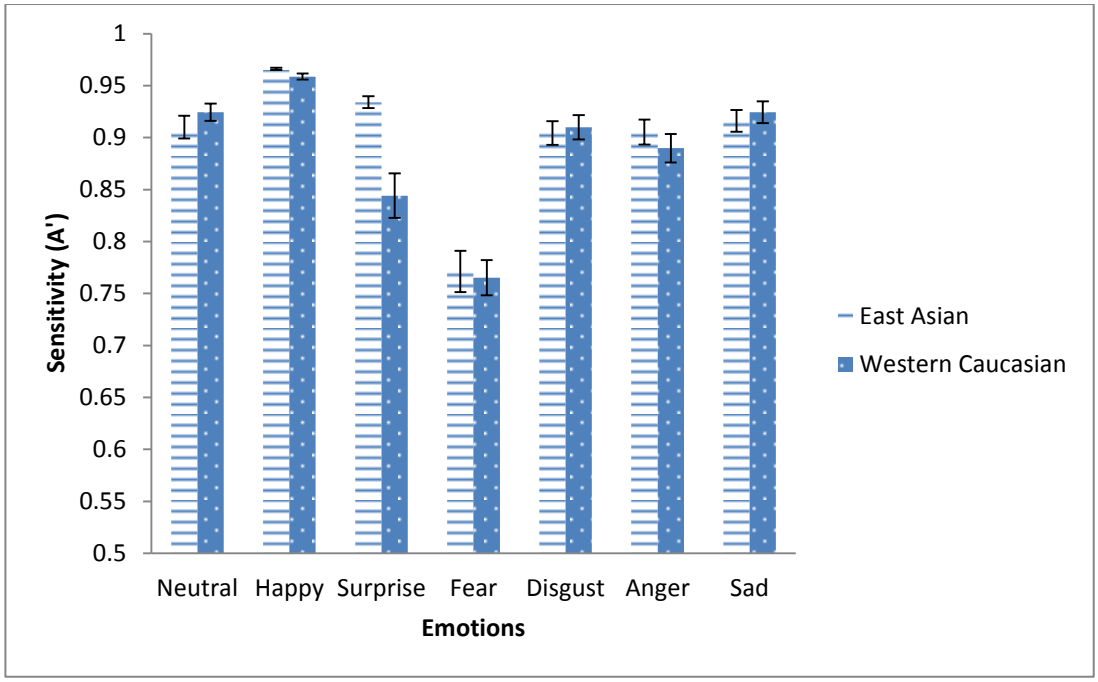


Figure 4.4. Malaysian Chinese participants' categorization sensitivity for East Asian and Western Caucasian emotions. Error bars report standard errors of the mean. Participants were best at categorizing "happy" while "fear" was categorized significantly worse than all other expressions. In addition, participants recognized East Asian expressions significantly more accurately than Western Caucasian ones.

In addition to the main effects, there was also an interaction between race of face and emotion, $F(3.68, 92.1)=6.01$, $p<.001$, Greenhouse-Geisser corrected. Post-hoc tests using paired samples t-test revealed that "happy," $t(25)=2.65$, $p<.05$ and "surprise," $t(25)=4.09$, $p<.001$ were categorized significantly better for East Asian ("happy" $M=.97$, $SD=.01$; "surprise" $M=.93$, $SD=.03$) than Western Caucasian ("happy" $M=.96$, $SD=.01$; "surprise" $M=.84$, $SD=.11$) faces, while the remaining emotions were categorized equally well across races, see Figure 4.4.

4.4.2. Eye Tracking Results

Linear Mixed Modeling analyses were conducted to test the effect of facial expressions of emotion ("neutral," "happy," "surprise," "fear," "disgust," "anger,"

and “sad”), race of face (East Asian or Western Caucasian) and feature of face (eyes, mouth, or nose) on the four dependent variables: total number of first fixations, total fixation count, total fixation duration, and mean fixation duration, see Table 4.2 for a summary of statistics for each dependent variable. Analyses for total fixation duration revealed a pattern similar to those observed in the total fixation count; therefore, the results for total fixation duration were not discussed in the current chapter (refer to Appendix B for the results of total fixation duration).

Table 4.2. Summary of statistics for each dependent variable in the linear mixed models.

	Total Number of First Fixations	**Total Fixation Count**	**Total Fixation Duration**	**Mean Fixation Duration**
Emotion	$F(6, 4158)=.01$, $p=1$	$F(6, 4152)=5.14$, $p<.001$**	$F(6, 4158)=2.82$, $p=.01$*	$F(6, 2640.28)=2.84$, $p=.009$*
Race	$F(1, 4158)=.01$, $p=.944$	$F(1, 6)=.01$, $p=.935$	$F(1, 4158)=.05$, $p=.825$	$F(1, 6.09)=.12$, $p=.745$
Feature	$F(2, 4158)=292.65$, $p<.001$**	$F(2, 4152)=610.57$, $p<.001$**	$F(2, 4158)=380.95$, $p<.001$**	$F(2, 3614.05)=41.71$, $p<.001$**
Emotion x Race	$F(6, 4158)=.01$, $p=1$	$F(6, 4152)=.93$, $p=.473$	$F(6, 4158)=.52$, $p=.794$	$F(6, 3640.28)=.87$, $p=.519$
Emotion x Feature	$F(12, 4158)=2.20$, $p=.009$*	$F(12, 4152)=5.14$, $p<.001$**	$F(12, 4158)=5.20$, $p<.001$**	$F(12, 3640.14)=1.06$, $p=.389$
Race x Feature	$F(2, 4158)=2.70$, $p=.067$	$F(2, 4152)=.47$, $p=.624$	$F(2, 4158)=1.08$, $p=.341$	$F(2, 3641.05)=1.31$, $p=.271$
Emotion x Race x Feature	$F(12, 4158)=.88$, $p=.567$	$F(12, 4152)=.94$, $p=.510$	$F(12, 4158)=2.01$, $p=.02$*	$F(12, 3640.14)=1.84$, $p=.037$*

* for $p<.05$

** for $p<.001$

4.4.2.1. Total Number of First Fixations

For total number of first fixations, there was a main effect of feature, $F(2, 4158)=292.64$, $p<.001$, an interaction between emotion and feature, $F(12, 4158)=2.20$, $p=.009$, and a trend towards an interaction between race and feature, $F(2, 4158)=2.70$, $p=.067$, see Figure 4.5. Post-hoc tests using Bonferroni Pairwise Comparisons showed that participants tended to direct their attention first to the nose more than the eyes ($p<.001$) and mouth ($p<.001$), and the eyes more than the mouth ($p<.001$).

To examine the interaction between emotion and feature, separate Linear Mixed Models were conducted for each of the seven emotions. Participant ID, face ID, and feature were included as factors (dependent variable = total number of first fixations; fixed factor = feature of face). There was a main effect of feature for all seven emotions ($p<.001$ for all emotions). Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for "neutral" faces, participants first fixated on the nose ($p<.001$) and eyes ($p<.001$) more than the mouth; however, the total number of first fixations landing on the eyes and nose did not differ significantly ($p=.41$). For all other emotions, participants tended to first fixate on the nose more than the eyes ($p<.001$ for "happy," "surprise," "fear," "disgust," "anger," and "sad" faces) and mouth ($p<.001$ for "happy," "surprise," "fear," "disgust," "anger," and "sad" faces), and eyes more than the mouth ($p<.001$ for "surprise," "fear," "anger," and "sad" faces; $p=.004$ for "happy" faces; $p=.029$ for "disgust" faces). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and emotion were included as factors

(dependent variable = total number of first fixations; fixed factor = emotion). There was a main effect of emotion for the mouth, $F(6, 1393)= 3.05$, $p=.006$, a trend towards a main effect for the eyes, $F(6, 1393)= 1.86$, $p=.085$, but no main effect for the nose, $F(6, 1393)= .45$, $p=.848$. Post-hoc tests using Bonferroni Pairwise Comparisons showed that there were significantly more first fixations landing on the mouth for "happy" ($p=.018$) and "disgust" ($p=.03$) than "sad" faces. As for the eyes, more first fixations were directed at "neutral" than "disgust" faces ($p=.048$).

To examine the trend towards an interaction between race and feature, separate Linear Mixed Models were conducted for each of the two races. Participant ID, face ID, and feature were included as factors (dependent variable = total number of first fixations; fixed factor = feature of face). There was a main effect of feature for East Asian, $F(2, 2097)=148.17$, $p<.001$ and Western Caucasian faces, $F(2, 2097)=147.10$, $p<.001$. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for both races of faces, participants tended to direct their attention first to the nose significantly more than the eyes ($p<.001$ for both races of faces) and mouth ($p<.001$ for both races of faces), and the eyes more than the mouth ($p<.001$ for both races of faces). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and race were included as factors (dependent variable = total number of first fixations; fixed factor = race of face). There was a trend towards a main effect of race for eyes, $F(1, 1398)=3.24$, $p=.072$, but not the mouth, $F(1, 1398)=1.29$, $p=.257$ and nose, $F(1, 1398)=.73$, $p=.392$. For the eyes, there were more first fixations for East Asian than Western Caucasian faces ($p=.072$).

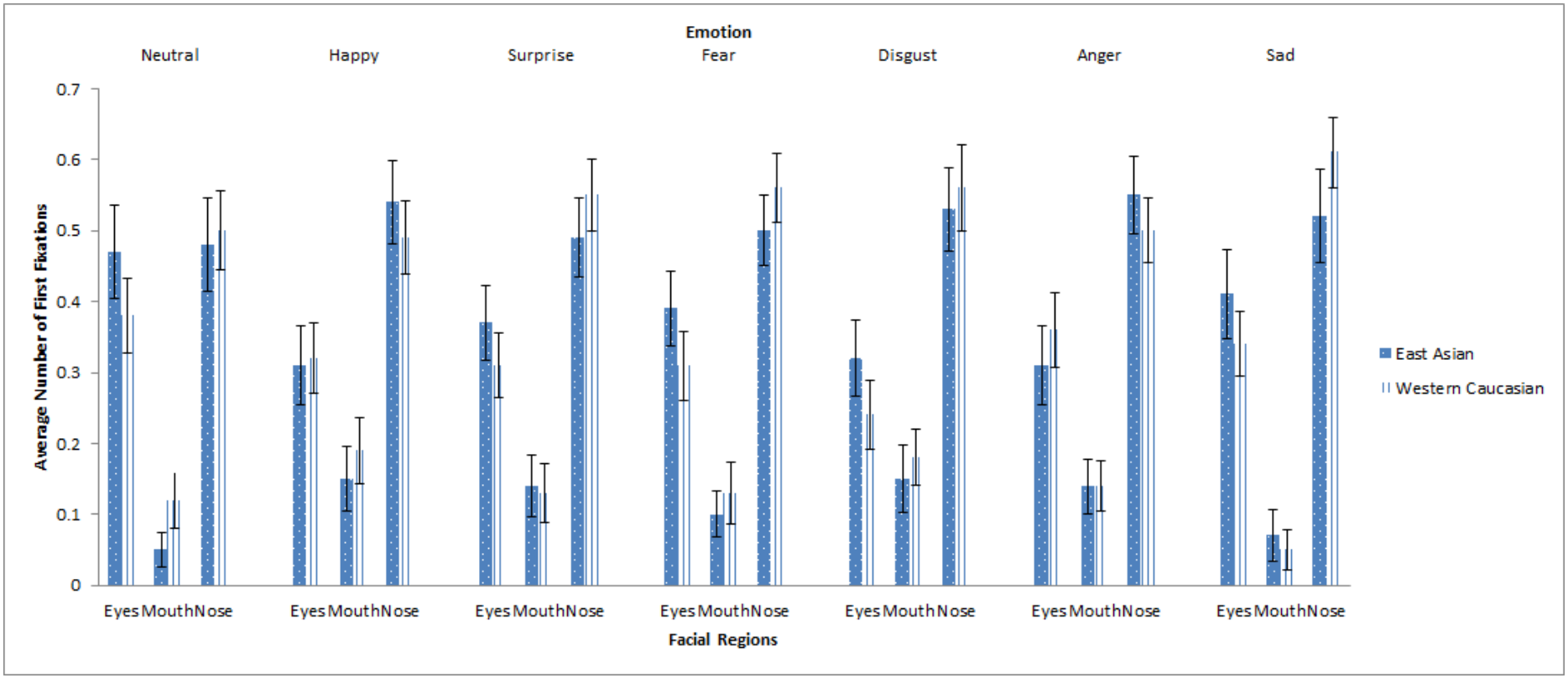


Figure 4.5. Average number of first fixations landing on the eyes, mouth, and nose during the emotion recognition task. Error bars report standard errors of mean. Participants directed their attention first to the nose more than the eyes and mouth, and eyes more than the mouth.

4.4.2.2. Total Fixation Count

For total fixation count, there was a main effect of emotion, $F(6, 4152)=5.14$, $p<.001$, and feature, $F(2, 4152)=610.56$, $p<.001$, and an interaction between emotion and feature, $F(12, 4152)=5.14$, $p<.001$, see Figure 4.6. For emotion, post-hoc tests using Bonferroni Pairwise Comparisons revealed that participants looked at "fearful" faces significantly more than "neutral" ($p=.032$) and "happy" ($p<.001$) faces, and a trend towards more fixations for "fearful" than "surprised" faces ($p=.052$). There were also more fixations directed at "angry" than "happy" faces ($p=.018$) and a trend towards more fixations for "sad" than "happy" faces ($p=.081$). As for feature, post-hoc tests using Bonferroni Pairwise Comparisons showed that participants fixated on the eyes significantly more frequently than the nose ($p<.001$) and mouth ($p<.001$), and the nose significantly more than the mouth ($p<.001$).

To examine the interaction between emotion and feature, separate Linear Mixed Models were conducted for each of the seven emotions. Participant ID, face ID, and feature were included as factors (dependent variable = total fixation count; fixed factor = feature of face). There was a main effect of feature for all seven emotions, $F(2, 597)=132.94$, $p<.001$ for "neutral," $F(2, 597)=50.38$, $p<.001$ for "happy," $F(2, 597)=56.47$, $p<.001$ for "surprise," $F(2, 597)=101.46$, $p<.001$ for "fear," $F(2, 597)=86.67$, $p<.001$ for "disgust," $F(2, 597)=105.96$, $p<.001$ for "anger," and $F(2, 597)=105.60$, $p<.001$ for "sad" faces. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for "neutral," "surprise," "fear," and "anger" faces, participants

made significantly more fixations on the eyes than the nose (p<.001 for "neutral" and "fear;" p=.004 for "surprise" and p=.003 for "anger") and mouth (p<.001 for all four emotions), and the nose more than the mouth (p<.001 for all four emotions). For "happy," "disgust," and "sad" faces, participants tended to fixate more on the eyes (p<.001 for all three emotions) and nose (p<.001 for all three emotions) than the mouth; however the total number of fixations for the eyes and nose did not differ significantly (p=.65 for "happy" faces, p=.679 for "disgust" faces, and p=.376 for "sad" faces). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and emotion were included as factors (dependent variable = total fixation count; fixed factor = emotion). There was a main effect of emotion for the eyes, $F(6, 1393)= 5.70$, p<.001, mouth, $F(6, 1393)=4.95$, p<.001, and nose $F(6, 1393)=4.15$, p<.001). Post-hoc tests using Bonferroni Pairwise Comparisons showed that there were significantly more fixations for the eyes for "neutral" (p=.003), "fear" (p<.001), and "anger" (p=.01) than "happy" faces, more fixations for "fear" and "neutral" than "disgust" faces (p=.002), and a trend towards more fixations for "fear" than "surprise" faces (p=.083) and "neutral" than "disgust" faces (p=.087). For the mouth, there were significantly more fixations for "surprise" than "neutral" (p<.001) and "anger" (p=.012) faces, more fixations for fear" than "neutral" (p=.003) faces, and a trend towards more fixations for the mouth for "surprise" than "sad" (p=.059) faces and "fear" than "anger" faces (p=.089). As for the nose, participants fixated significantly more on "disgust" than "neutral" (p=.033), "happy" (p=.003), and "surprise" (p=.019) faces, and a trend towards more fixation counts on "sad" than "happy" faces (p=.075).

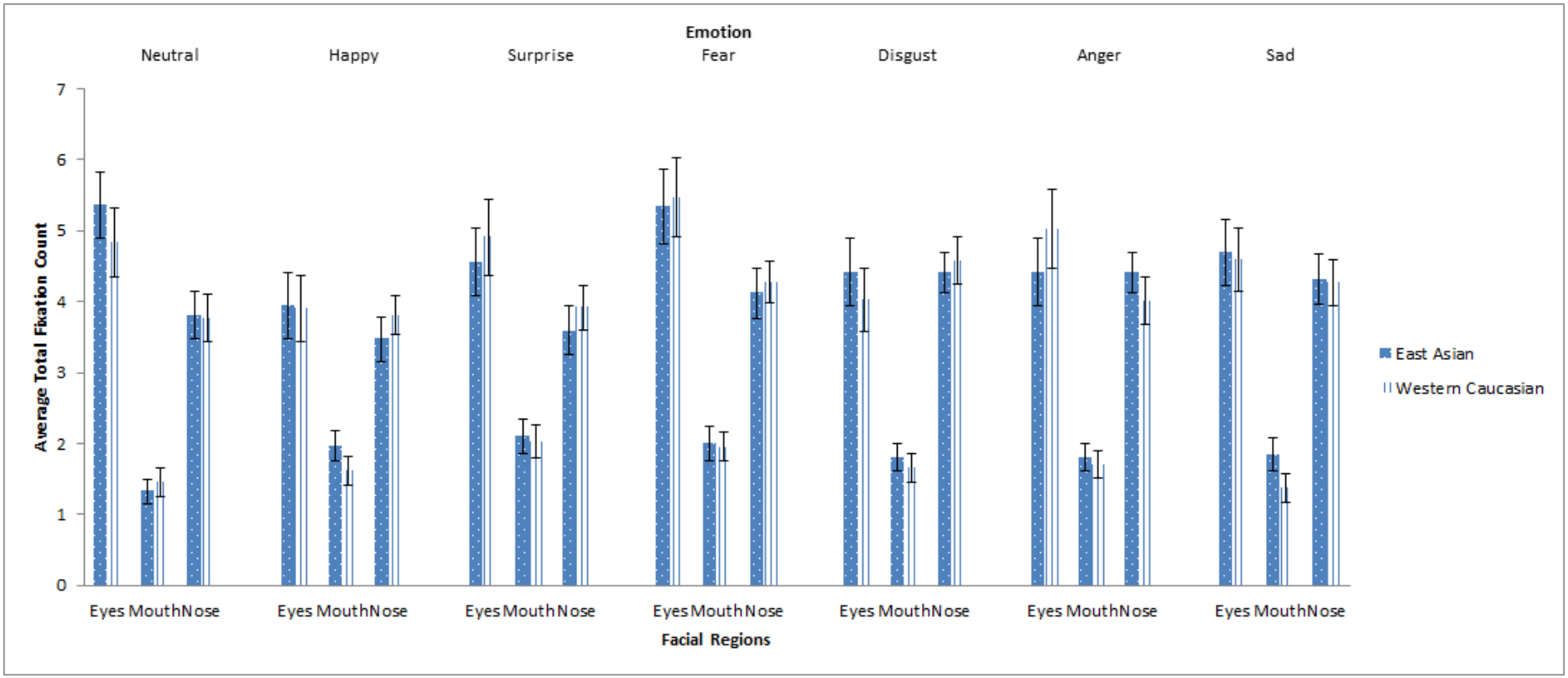


Figure 4.6. Average total fixation count for the eyes, mouth, and nose during the emotion recognition task. Error bars report standard errors of mean.

Participants fixated most frequently on the eyes, followed by the nose then mouth.

4.4.2.3. Mean Fixation Duration

For mean fixation duration, there was a main effect of emotion, $F(6, 3640.28)=2.84$, $p=.009$, feature, $F(2, 3641.05)=41.705$, $p<.001$, an interaction between emotion and feature, $F(12, 3640.14)=1.06$, $p<.001$, and an interaction between race, emotion, and feature, $F(12, 3640.14)=1.84$, $p=.037$, see Figure 4.7. For emotion, post-hoc tests using Bonferroni Pairwise Comparisons revealed that the mean fixation duration for "neutral" faces was significantly longer than "fearful" ($p=.003$) and "angry" ($p=.043$) faces. As for feature, post-hoc tests using Bonferroni Pairwise Comparisons showed that the mean fixation duration for the mouth was significantly longer than the eyes ($p<.001$) and nose ($p<.001$). However, the mean fixation duration for the eyes and nose did not differ significantly ($p=.083$).

To examine the interaction between emotion and feature, separate Linear Mixed Models were conducted for each of the seven emotions. Participant ID, face ID, and feature were included as factors (dependent variable = total fixation duration; fixed factor = feature of face). There was a main effect of feature for all seven emotions, $F(2, 511)=3.82$, $p=.023$ for "neutral," $F(2, 520)=5.78$, $p=.003$ for "happy," $F(2, 534)=16.63$, $p<.001$ for "surprise," $F(2, 539)=5.84$, $p=.003$ for "fear," $F(2, 528)=3.77$, $p=.024$ for "disgust," $F(2, 516)=10.00$, $p<.001$ for "anger," and $F(2, 518)=9.23$, $p<.001$ for "sad" faces. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for "surprise," "anger," and "sad" faces, the mean fixation duration for the mouth was significantly longer than the eyes ($p<.001$ for "surprise," $p=.001$ for "anger" and "sad") and nose ($p<.001$ for "surprise,"

"anger," and "sad" faces); however, the mean fixation duration for the eyes and nose did not differ significantly (p=.447 for "surprise," p=1 for "anger" and "sad" faces). For "fearful" faces, the mean fixation duration for the mouth (p=.004) and eyes (p=.032) was longer than the nose; however, the mean fixation duration for the mouth and eyes did not differ significantly (p=1). For "happy" and "disgust" faces, the mean fixation duration for the mouth was significantly longer than the nose (p=.002 for "happy" and p=.032 for "disgust" faces); however, the mean fixation duration for the mouth was not significantly longer than the eyes (p=.404 for "happy" and p=.08 for "disgust" faces) and the eyes was not significantly longer than the nose (p=.166 for "happy and p=1 for "disgust" faces). For "neutral" faces, the mean fixation duration for the mouth was significantly longer than the eyes (p=.031); however, the mean fixation duration for the mouth is not significantly different from the nose (p=.061), and the nose was not significantly longer than the eyes (p=1). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and emotion were included as factors (dependent variable = total fixation duration; fixed factor = emotion). There was a trend towards a main effect of emotion for the mouth, F(6, 1030)=1.92, p=.075, but not for the eyes, F(6, 1287)=.98, p=.435, and nose, F(6, 1349)=1.44, p=.196. Post-hoc tests using Bonferroni Pairwise Comparisons showed that the mean fixation duration for the mouth was significantly longer for "neutral" than "fearful" faces (p=.038).

To examine the interaction between race, emotion, and feature, two separate Linear Mixed Models were conducted for each race. For both East Asian and Western Caucasian faces, there was a main effect of feature only,

$F(2, 1832)=30.44$, $p<.001$ for East Asian faces and $F(2, 1813)=13.72$, $p<.001$ for Western Caucasian faces. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for both East Asian and Western Caucasian faces, the mean fixation duration for the mouth was significantly longer than the eyes ($p<.001$ for East Asian faces and $p=.001$ for Western Caucasian faces) and nose ($p<.001$ for both East Asian and Western Caucasian faces); however, the mean fixation duration for the eyes and nose did not differ significantly ($p=.403$ for East Asian faces and $p=.333$ for Western Caucasian faces).

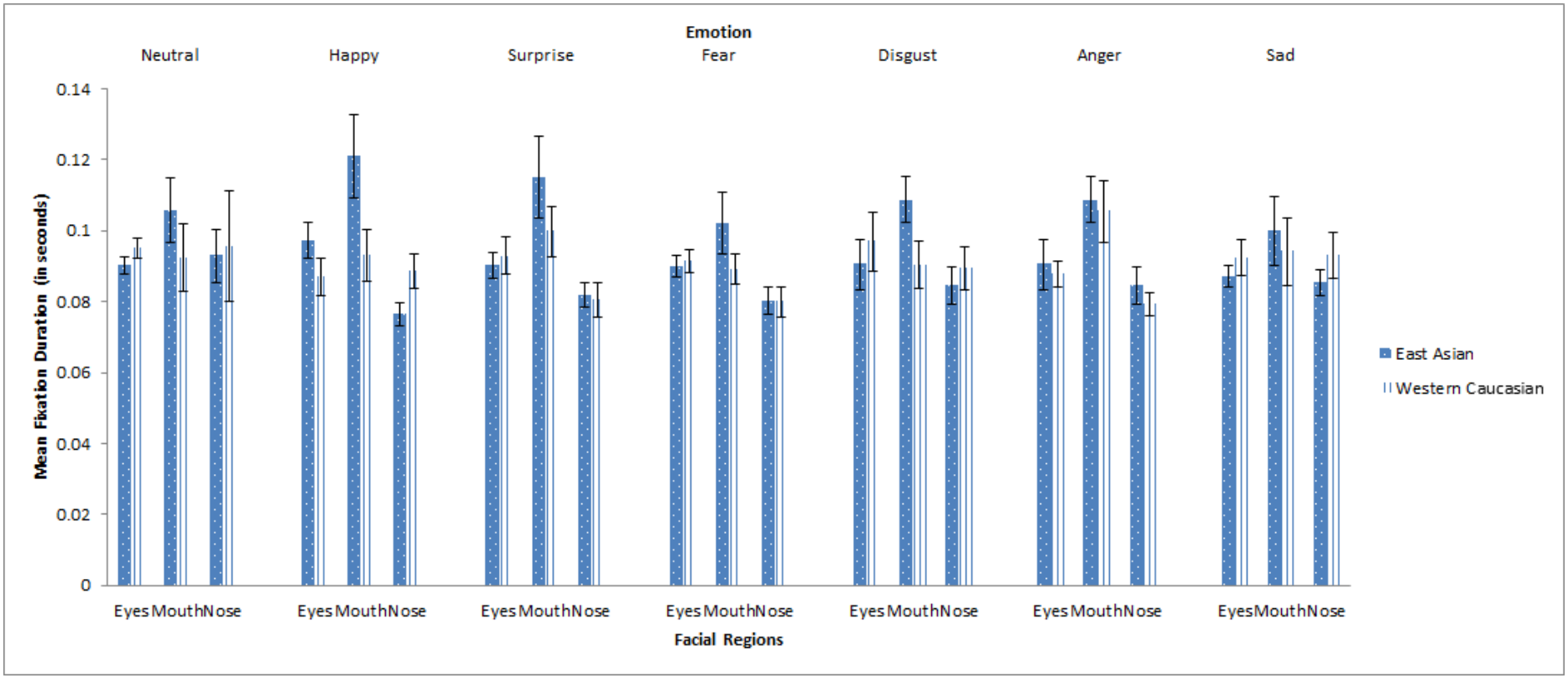


Figure 4.7. Mean fixation duration for the eyes, mouth, and nose during the emotion recognition task, in seconds. Error bars report standard errors of mean. The mean fixation duration for the mouth was significantly longer than the eyes and nose.

4.5. Static Face Recognition and Emotion Recognition Tasks

Tracking an observer's eye movements provides vital information about the elements that attract one's attention. Early research by Yarbus (1967) has shown that observers shift their visual attention to a particular element (e.g. features of a face or an object in a scene) only when it is believed to be useful and crucial for perception. Because the degree of attention given to an element is determined by the deemed importance of the element and the nature of the task, tracking an observer's eye movements would allow researchers to identify the elements that attract one's attention and provide insight into an observer's thought processes and cognitive goals (Yarbus, 1967; Hayhoe & Ballard, 2005; Andrews & Coppola, 1999).

The face processing research reviewed so far suggests that there are distinct Eastern and Western eye movement strategies employed in face recognition and emotion recognition tasks; strategies that may be affected by observers' goals and aims (Blais et al., 2008; Jack et al., 2009). To examine whether Malaysian Chinese participants change their eye movement strategies according to the nature of the task (i.e. face recognition and emotion recognition tasks), further analyses were conducted to identify the interaction between Malaysian Chinese participants' looking patterns with factors such as task (i.e. the recognition phase from the face recognition task and the emotion recognition task) and race of face (i.e. East Asian and Western Caucasian). The procedure for the face recognition task (as presented in Chapter Three) was similar to task described here, except that participants were asked to identify if they had previously seen a series of African, East Asian and Western Caucasian faces presented on the screen of a Tobii T60 eye tracker for 5 seconds each, instead of being asked to identify the emotional facial expression. Eye tracking data

for African faces in the face recognition task were not included in the current analyses and eye tracking data were collapsed across emotions in the emotion recognition task in order to be consistent.

Linear Mixed Modeling analyses were conducted to test the effect of task (static face recognition or emotion recognition tasks), race of face (East Asian or Western Caucasian) and feature of face (eyes, mouth, or nose) on the four dependent variables: total number of first fixations, total fixation count, total fixation duration, and mean fixation duration, see Table 4.3 for a summary of statistics for each dependent variable. Participant ID, face ID, task, race of face, feature of face were included as factors. All main effects and the two- and three-way interactions between task, race of face, and feature of face were included in the model (dependent variable = total number of first fixations, total fixation count, total fixation duration and mean fixation duration; fixed factors = task, race of face, and feature of face). Participant ID nested within task and face ID nested within race were included as a random factor for total fixation count and mean fixation duration. Only participant ID nested with task was included as a random factor for total fixation duration and no random factors were included for total number of first fixations because the final Hessian matrix was not positive, which casts doubt on the validity of results; however, all results with or without face ID nested within race included as a random factor were identical.

Analyses for total fixation duration revealed a pattern similar to those observed in the total fixation count; therefore, the results for total fixation duration were not discussed in the current chapter (refer to Appendix C for the results of total fixation duration).

Table 4.3. Summary of statistics for each dependent variable in the linear mixed models.

	Total Number of First Fixations	**Total Fixation Count**	**Total Fixation Duration**	**Mean Fixation Duration**
Task	$F(1, 6588)=.004$, $p=.948$	$F(1, 43.96)=4.42$, $p=.041$*	$F(1, 42.73)=22.50$, $p<.001$**	$F(1, 45.10)=6.11$, $p=.017$*
Race	$F(1, 6588)=.00$, $p=.994$	$F(1, 41.64)=.33$, $p=.570$	$F(1, 6543.84)=.03$, $p=.860$	$F(1, 34.27)=.37$, $p=.550$
Feature	$F(2, 6588)=413.89$, $p<.001$**	$F(2, 6532.05)=956.73$, $p<.001$**	$F(2, 6543.81)=638.68$, $p<.001$**	$F(2, 5742.57)=37.27$, $p<.001$**
Task x Race	$F(1, 6588)=.01$, $p=.939$	$F(1, 41.64)=.50$, $p=.482$	$F(1, 6543.84)=.21$, $p=.645$	$F(1, 34.27)=.01$, $p=.944$
Task x Feature	$F(2, 6588)=14.07$, $p<.001$*	$F(2, 6532.05)=2.99$, $p=.050$*	$F(2, 6543.81)=3.27$, $p=.038$*	$F(2, 5742.57)=5.81$, $p=.003$*
Race x Feature	$F(2, 6588)=.60$, $p=.549$	$F(2, 6532.06)=1.48$, $p=.227$	$F(2, 6543.84)=2.74$, $p=.064$	$F(2, 5732.34)=2.68$, $p=.069$
Task x Race x Feature	$F(2, 6588)=4.16$, $p=.016$*	$F(2, 6532.06)=.55$, $p=.578$	$F(2, 6543.84)=1.31$, $p=.271$	$F(2, 5732.34)=.11$, $p=.896$

* for $p<.05$

** for $p<.001$

4.5.1. Total Number of First Fixations

For total number of first fixations, there was a main effect of feature, $F(2, 6588)=413.89$, $p<.001$, an interaction between task and feature, $F(2, 6588)=14.07$, $p<.001$, and an interaction between task, race, and feature, $F(2, 6588)=4.16$, $p=.016$, see Figure 4.8. Post-hoc tests using Bonferroni Pairwise Comparisons showed that participants tended to direct their attention first to the nose more than the eyes ($p<.001$) and mouth ($p<.001$), and the eyes more than the mouth ($p<.001$).

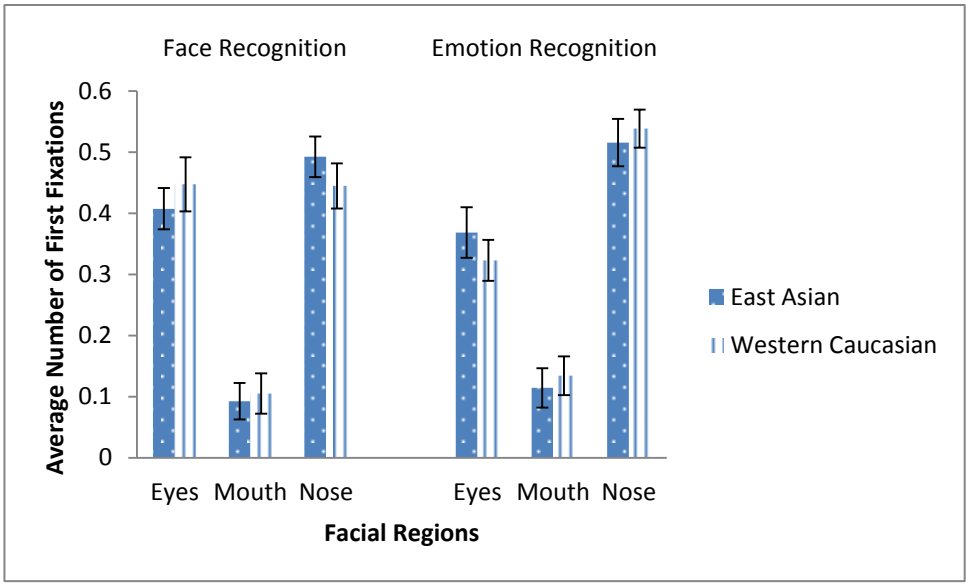


Figure 4.8. Average number of first fixations landing on the eyes, mouth, and nose during the face recognition and emotion recognition tasks. Error bars report standard errors of mean. For the face recognition task, participants first fixated on the nose and eyes more than the mouth, while for the emotion recognition task, there were more first fixations for the nose, followed by the eyes then mouth.

To examine the interaction between task and feature, separate Linear Mixed Models were conducted for each of the two tasks. Participant ID, face ID, and feature were included as factors (dependent variable = total number of first fixations; fixed factor = feature of face). There was a main effect of feature for both face recognition

and emotion recognition task, $F(2, 2397)=169.10$, $p<.001$ for the face recognition task and $F(2, 4197)=292.40$, $p<.001$ for the emotion recognition task. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for the face recognition task, participants first fixated on the nose ($p<.001$) and eyes ($p<.001$) more than the mouth; however, the total number of first fixations landing on the eyes and nose did not differ ($p=.185$). For the emotion recognition task, participants tended to first fixate on the nose more than the eyes ($p<.001$) and mouth ($p<.001$), and eyes more than the mouth ($p<.001$). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and task were included as factors (dependent variable = total number of first fixations; fixed factor = task). There was a main effect of task for the eyes, $F(1, 2198)=14.61$, $p<.001$, nose, $F(1, 2198)=6.96$, $p=.008$, and a trend for the mouth, $F(1, 2198)=3.26$, $p=.071$. For the eyes and nose, there were significantly more first fixations for the face recognition than emotion recognition task ($p<.001$ for the eyes and $p=.008$ for the nose), whereas for the mouth, there was a trend towards more first fixations on the emotion recognition than face recognition task ($p=.07$).

To examine the interaction between task, race, and feature, two separate Linear Mixed Models were conducted for each task. For the face recognition task, there was a main effect of feature only, $F(2, 2394)=169.17$, $p<.001$. For feature, post-hoc tests using Bonferroni Pairwise Comparisons revealed that for the face recognition task, participants tended to fixate first on the nose ($p<.001$) and eyes ($p<.001$) significantly more than the mouth; however the total number of first fixations landing on the nose and eyes did not differ significantly ($p=.186$).

As for the emotion recognition task, there was a main effect of feature, $F(2, 4194)=292.56$, $p<.001$, and a trend towards an interaction between race and feature,

$F(2, 4194)=2.70$, $p=.067$. For feature, post-hoc tests using Bonferroni Pairwise Comparisons revealed that for the emotion recognition task, participants tended to fixate first on the nose significantly more than the eyes ($p<.001$) and mouth ($p<.001$), and the eyes significantly more than the mouth ($p<.001$). To examine the trend towards an interaction between race and feature for the emotion recognition task, separate Linear Mixed Models were conducted for each of the two races. Participant ID, face ID, and feature were included as factors (dependent variable = total number of first fixations; fixed factor = feature of face). There was a main effect of feature for both races, $F(2, 2097)=138.17$, $p<.001$ for East Asian faces and $F(2, 2097)=147.10$, $p<.001$ for Western Caucasian faces. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for both East Asian and Western Caucasian faces, the total number of first fixations landing on the nose was significantly more than the eyes ($p<.001$) and mouth ($p<.001$), and the eyes more than the mouth ($p<.001$). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and race were included as factors (dependent variable = total number of first fixations; fixed factor = race of face). There was a trend towards a main effect of race for the eyes, $F(1, 1398)=3.24$, $p=.072$, but there was no main effect for the mouth, $F(1, 1398)=1.29$, $p=.257$, and nose, $F(1, 1398)=.73$, $p=.392$. For the eyes, the total number of first fixations landing on East Asian faces was significantly more than Western Caucasian faces ($p=.072$).

4.5.2. Total Fixation Count

For total fixation count, there was a main effect of task, $F(1, 43.96)=4.412$, $p=.041$, and feature, $F(2, 6532.05)=956.73$, $p<.001$, and an interaction between task and feature, $F(2, 6532.05)=2.99$, $p=.05$, see Figure 4.9. For task, there were more

fixations for the emotion recognition than face recognition task (p=.041). As for feature, post-hoc tests using Bonferroni Pairwise Comparisons showed that participants fixated on the eyes significantly more than the nose (p<.001) and mouth (p<.001), and the nose significantly more than the mouth (p<.001).

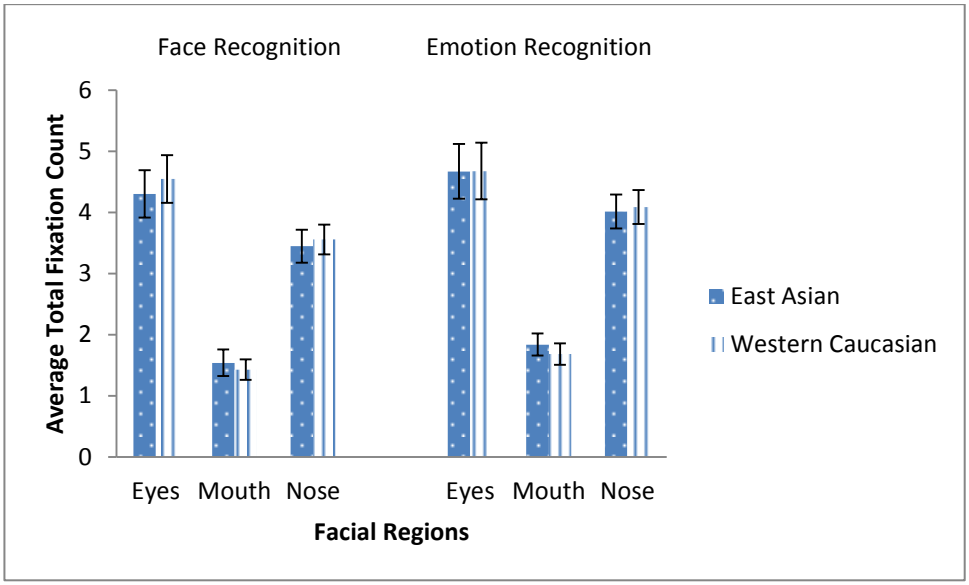


Figure 4.9. Average total fixation count for the eyes, mouth, and nose during the face recognition and emotion recognition tasks. Error bars report standard errors of mean. Participants fixated most frequently on the eyes, followed by the nose then mouth for both tasks.

To examine the interaction between task and feature, separate Linear Mixed Models were conducted for each of the two tasks. Participant ID, face ID, and feature were included as factors (dependent variable = total fixation count; fixed factor = feature of face). There was a main effect of feature for both face recognition and emotion recognition tasks, F(2, 2397)=407.68, p<.001 for face recognition task and F(2, 4197)=599.57, p<.001 for emotion recognition task. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for both tasks, participants made significantly more fixations on the eyes than the nose (p<.001 for both tasks) and

mouth (p<.001 for both tasks), and the nose more than the mouth (p<.001 for both tasks). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and task were included as factors (dependent variable = total fixation count; fixed factor = task). There was a main effect of task for all three features, with high significance level for the nose, $F(1, 2198)=31.85$, p<.001, and mouth, $F(1, 2198)=13.74$, p<.001, and a marginal significance for the eyes, $F(1, 2198)=3.92$, p=.048. Although participants fixated more on the emotion than face recognition task for all three features, participants tended to shift their attention away from the eyes and towards the nose and mouth in the emotion recognition task (p<.001 for the nose and mouth and p=.048 for the eyes), see Figure 4.10.

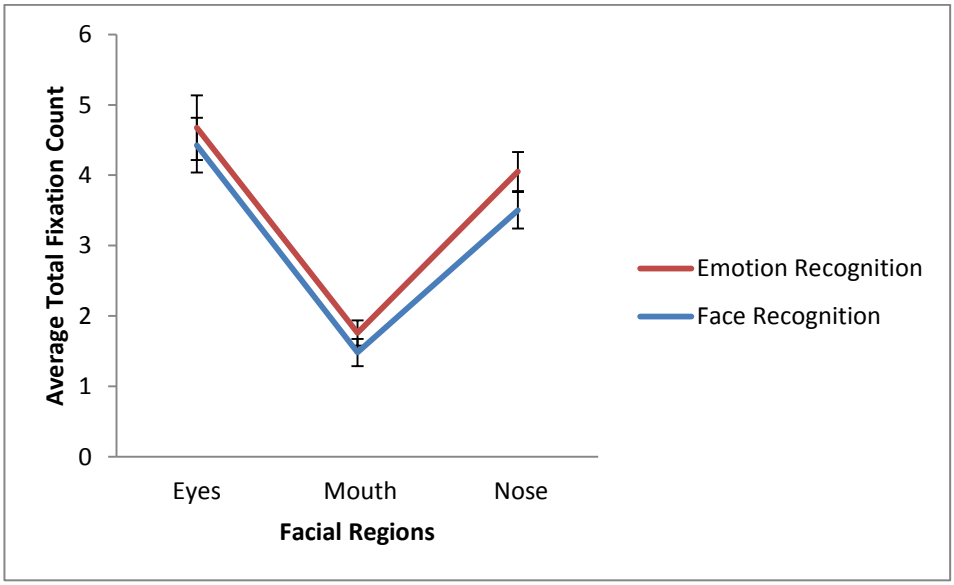


Figure 4.10. Average total fixation count for the eyes, mouth, and nose during the emotion recognition and face recognition tasks. Participants shifted their attention away from the eyes and towards the nose and mouth in the emotion recognition task.

4.5.4. Mean Fixation Duration

For mean fixation duration, there was a main effect of task, $F(1, 45.10)=6.11$, p=.017, feature, $F(2, 5742.57)=37.27$, p<.001, an interaction between task and feature,

$F(2, 5742.57)=5.81$, $p=.003$, and a trend towards an interaction between race and feature, $F(2, 5732.34)=2.68$, $p=.069$, see Figure 4.11. For task, the mean fixation duration for the emotion recognition task was significantly longer than the face recognition task ($p=.017$). As for feature, post-hoc tests using Bonferroni Pairwise Comparisons showed that the mean fixation duration for the mouth was significantly longer than the eyes ($p<.001$) and nose ($p<.001$), and the eyes longer than the nose ($p=.02$).

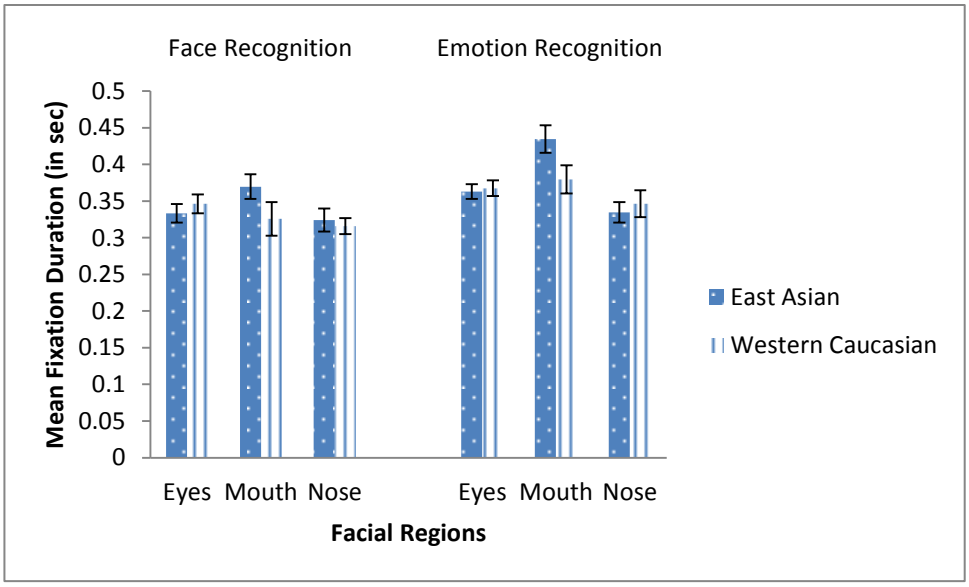


Figure 4.11. Mean fixation duration for the eyes, mouth, and nose during the face recognition and emotion recognition tasks, in seconds. Error bars report standard errors of mean. For the face recognition task, the mean fixation duration for the mouth and eyes was significantly longer than the nose, whereas for the emotion recognition task, the mean fixation duration for the mouth was significantly longer than the eyes and nose.

To examine the interaction between task and feature, separate Linear Mixed Models were conducted for each of the two tasks. Participant ID, face ID, and feature were included as factors (dependent variable = mean fixation duration; fixed factor = feature of face). There was a main effect of feature for both face recognition and

emotion recognition tasks, $F(2, 2107)=10.46$, $p<.001$ for face recognition task and $F(2, 3684)=41.85$, $p<.001$ for emotion recognition task. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for the face recognition task, the mean fixation duration for the mouth ($p<.001$) and eyes ($p=.004$) was significantly longer than the nose; however, the mean fixation duration for the mouth and eyes did not differ significantly ($p=.453$). For the emotion recognition task, the mean fixation duration for the mouth was significantly longer than the eyes ($p<.001$) and nose ($p<.001$); however the mean fixation duration for the eyes and nose did not differ significantly ($p=.092$). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and task were included as factors (dependent variable = mean fixation duration; fixed factor = task). There was a main effect of emotion for the eyes, $F(1, 2059)=7.40$, $p=.007$, mouth, $F(1, 1618)=36.36$, $p<.001$, and nose, $F=(1, 2114)=7.91$, $p=.005$). For the eyes, mouth, and nose, the mean fixation duration was significantly longer for the emotion recognition than face recognition task ($p=.007$ for the eyes, $p<.001$ for the mouth, and $p=.005$ for the nose).

To examine the trend towards an interaction between race and feature, separate Linear Mixed Models were conducted for each of the two races. Participant ID, face ID, and feature were included as factors (dependent variable = mean fixation duration; fixed factor = feature of face). There was a main effect of feature for both races, $F(2, 2904)=37.35$, $p<.001$ for East Asian faces and $F(2, 2887)=15.77$, $p<.001$ for Western Caucasian faces. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for East Asian faces, the mean fixation duration for the mouth is significantly longer than the eyes ($p<.001$) and nose ($p<.001$); however, the mean fixation duration for the eyes and nose did not differ significantly ($p=.13$). For Western Caucasian faces, the mean fixation duration for the mouth was significantly

longer than the eyes (p=.007) and nose (p<.001), and the eyes longer than the nose (p=.019). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and race were included as factors (dependent variable = mean fixation duration; fixed factor = race of face). There was a trend towards a main effect of race for the mouth, $F(1, 1618)=3.79$, p=.052, but not for the eyes, $F(1, 2059)=1.44$, p=.23, and nose, $F(1, 2114)=.024$, p=.876. For the mouth, the mean fixation duration for East Asian faces was significantly longer than Western Caucasian faces (p=.052).

4.6. Discussion

The current study explored Malaysian Chinese participants' eye movement strategy and categorization sensitivity in recognizing East Asian and Western Caucasian facial expressions of emotion. The primary findings were that 1) Malaysian Chinese participants were best at categorizing "happy" while "fear" was recognized less well than other emotions, 2) attention was directed first to the nose more than the eyes, followed by the mouth, 3) participants fixated most frequently on the eyes, followed by the nose then mouth, and 4) the mean fixation duration for the mouth was significantly longer than the eyes and nose.

Behavioral results showed that Malaysian Chinese participants were best at categorizing "happy" whereas "fear" was recognized less well than other emotions. Additionally, Malaysian Chinese were able to recognize East Asian expressions better than Western Caucasian ones, although this was only true for "happy" and "surprise." It is unlikely that this is due to unfamiliarity with Western Caucasian faces, because in the face recognition study (presented in Chapter Three), it was shown that Malaysian Chinese recognized Western Caucasian faces with equal proficiency as East Asian

ones. A more plausible argument might be that due to different cultural interpretation of certain emotions (Ekman & Friesan, 1971), for other-race faces, there is a discrepancy between the way expressions were posed and perceivers' conception of those emotions. Observers may find certain portrayal of emotions more relevant or familiar to their culture, and hence be more proficient in recognizing these expressions. In any case, Malaysian Chinese observers' equal success in recognizing East Asian and Western Caucasian identities does not appear to extend to interpreting their facial expressions.

"Fear" was identified significantly less well than other emotions. Half of the participants mistook "fear" as "surprise," a finding similar to Ekman and Friesman (1971)'s study with New Guineans. Ekman and Friesman (1971) speculated that participants confused the two expressions, not because of their inability to distinguish between the two, but because fearful situations are almost always regarded as surprising in their culture. In addition, a cross-cultural study revealed that although observers from different cultures agreed on the emotions portrayed in each photograph, the levels of agreement between cultures were considerably different (Biehl, Matsumoto, Ekman, Hearn, Heider, Kudoh, & Ton, 1997). Japanese observers, for instance, had less agreement on negative expressions like "fear" as compared to other countries whereas Americans agreed less than other countries on "contempt," suggesting that the semantic or affect meanings of each emotion may vary across culture. Observers also rated the intensity of emotion differently, indicating dissimilarities in either the portrayal or judgment of facial expressions. Therefore, it is probable that Malaysian Chinese had difficulty recognizing "fear" due to the different meaning or judgment of the emotion.

When categorizing facial expressions, Malaysian Chinese observers tended to direct their attention first to the nose and eyes more than the mouth for "neutral" faces whereas for all other emotions, participants first fixated on the nose more than the eyes and mouth. In addition, there were more first fixations landing on the mouth for "happy" and "disgust" than "sad" faces. Participants tended to fixate first on the center of the face (in the nose area) to gain a more holistic view of the face at first glance possibly because participants are uncertain which facial expressions will be displayed on screen and the expressive region of the face may vary according to the emotions; for instance, "happy" and "disgust" faces may be more expressive in the mouth region as compared to "sad" faces.

Malaysian Chinese observers fixated mostly on the eyes, followed by the nose then mouth when recognizing "neutral," "surprise," "fear," and "anger" faces while there were more fixations on the eyes and nose than the mouth for "happy," "disgust," and "sad" faces. Participants had a greater preference for the eyes for emotions like "neutral," fear," and "anger," a stronger preference for the mouth for "surprise" and "fear" faces, while participants preferred the nose for "disgust" faces. Participants' fixation pattern varied according to the facial expressions displayed, suggesting that participants altered their eye movement strategy to focus on the expressive regions of the face, which may differ according to the emotion being portrayed. This finding is similar to previous eye tracking results which also found that participants attended to different regions of the face to identify different facial expressions (Guo, 2012).

Malaysian Chinese participants' eye movement strategy employed in the current static emotion recognition task (i.e. more fixations on the eyes, followed by the nose then mouth) appear similar to Western Caucasians' eye movement strategy reported in Guo's (2012) study, but was different from those reported in Jack's (2009)

study, which found that Western Caucasians made even distribution of fixations on the eyes and mouth. Comparison of the eye tracking results between the current study and Guo's (2012) study suggest that the eye movements of Malaysian Chinese and Western Caucasian observers may not be so different; however, the results found in both studies may not be directly comparable as the methodologies employed in both studies are different (i.e. the current study required Malaysian Chinese participants to categorize own- and other-race facial expressions of emotion while Guo's (2012) study required Western Caucasian participants to categorize own-race facial expressions morphed to 5 levels of intensity). Guo (2012) suggested that facial expressions with lower intensities may be more ambiguous; therefore, Western Caucasian participants may fixate on the nose to gain a more holistic view of the face as the cues in the eyes and mouth may not be as evident in expressions with lower intensities. The eye tracking results of the current study may be more comparable to Jack's (2009) study, which found that Western Caucasians moved between the eyes and mouth while East Asians focused on the eye region alone when categorizing emotions, as both studies employed similar methodologies (Jack et al., 2009).

Lastly, the mean fixation duration showed that participants made fixations of a longer duration on the mouth than the eyes and nose, implying that participants made longer but fewer fixations on the mouth and shorter but more fixations on the eyes and nose to identify emotions. Additionally, the mean fixation duration also showed that participants focused more on different parts of the face depending on the facial expressions presented, a pattern similar to those found in the total fixation count and in previous eye tracking study (Guo, 2012). For "surprise," "anger," and "sad" faces, the mean fixation duration for the mouth was longer than the eyes and nose; for "fearful" faces, the mean fixation duration for the mouth and eyes was longer than the

nose; for “happy” and “disgust” faces, the mean fixation duration for the mouth was longer nose, but the mean fixation duration for the eyes was not different from those of the mouth and nose; as for “neutral” faces, the mean fixation duration for the mouth was longer than the eyes, but the mean fixation duration for the nose was not different from those of the mouth and eyes. In addition, the mean fixation duration for the mouth was longer for “neutral” than “fearful” faces.

Recent cross-cultural face processing studies have found that East Asian observers tended to focus in the nose area when performing a face recognition task; however, observers focused on the eye region alone when asked to identify emotion rather than identity, suggesting an adaptation of strategy (Blais et al., 2008; Jack et al., 2009). Observers may have learned to adopt a different fixation pattern (by directing their attention towards the eyes) when performing the emotion recognition task possibly because certain emotions are expressed differently in the East. Additionally, East Asian observers’ lack of attention to the mouth may have contributed to the misrecognition of certain emotions such as “fear” and “disgust,” which were consistently confused with “surprise” and “anger” (Jack et al., 2009). Conversely, Western Caucasian observers appear to employ a similar strategy for both face recognition and emotion recognition tasks by moving between the eyes and mouth to recognize identity and emotion (Blais et al., 2008; Jack et al., 2009). Western Caucasian observers may have employed a similar fixation pattern in the emotion recognition task possibly because fixating on the expressive regions of the face (i.e. the eyes and mouth) was an optimal strategy to recognize emotion as well as recognize identities. The even distribution of fixations across faces enabled Western Caucasians to recognize all facial expressions of emotion with relatively high accuracy. These findings suggest that observers employ different eye movement

strategies according to the nature of the task (Blais et al., 2008; Jack et al., 2009). Therefore, besides examining Malaysian Chinese observers' performance on the emotion recognition task, the current results were further compared with the face recognition data presented in Chapter Three, revealing interactions between task and race for certain regions of the face.

For the face recognition task, Malaysian Chinese participants tended to fixate first on the nose and eyes more than the mouth; however, there were more first fixations landing on the nose, followed by the eyes then mouth for the emotion recognition task. Fixating first on the nose is an optimal strategy for participants to gain a holistic view of the face; additionally, the center of the face is a strategically ideal position from which fixations can be swiftly shifted towards either the eyes or mouth when recognizing both identities and emotions (Vo, Smith, Mital, & Henderson, 2012). However, when performing the face recognition task, participants tended to fixate first on both the nose and the eyes possibly because the individuating feature of the face (i.e. the eyes) may provide important cues to recalling identities. Also, participants tended to direct their attention first to the eyes for East Asian faces significantly more than Western Caucasian faces when performing the emotion recognition task. Previous study has found that East Asian observers learned to rely on the eye region for more accurate information about East Asians' true emotion (Yuki et al., 2007). In line with this finding, Malaysian Chinese participants may fixate first on the eyes for East Asian faces more so than Western Caucasian ones as the eyes may contain diagnostic information relevant to East Asian expressions.

In both tasks, Malaysian Chinese appear to have employed a similar eye movement strategy by focusing on the eyes more than the nose, followed by the mouth when perceiving both East Asian and Western Caucasian faces. However,

analyses comparing the total fixation count landing on East Asian and Western Caucasian eyes, mouth, and nose for both tasks revealed that participants showed a tendency to look less at the eyes and more at the nose and mouth in the emotion recognition task, compared to the face recognition task. This change of strategy is consistent with previous research which found that East Asian observers use different eye movement patterns for these two types of face processing tasks; more specifically, East Asians fixated mostly on the eyes to recognize emotions because they bear diagnostic information of an individual's true emotion but employed a holistic strategy by focusing on the nose to recognize identities (Blais et al., 2008; Jack et al., 2009; Yuki et al., 2007). However, the nature of the change in strategy observed here for Malaysian Chinese participants was distinct from that observed in East Asian observers previously—attention was shifted towards the lower part of the face (i.e. the nose and mouth) for recognizing emotions than identities.

Tan et al. (2012) speculated that Malaysian Chinese participants adopted a combination of Eastern and Western eye movement strategies in the face recognition task to better recognize own- and other-race faces. Although Malaysian Chinese shifted their attention towards the nose and mouth in the emotion recognition task, suggesting a move towards a Western strategy, fixations were directed mainly towards the eyes, a strategy that may have enabled Malaysian Chinese to better recognize own- than other-race facial expressions. Previous studies have suggested that observers learn to rely on different facial cues because of the cultural differences in how the emotions are expressed (Yuki et al., 2007). East Asians tended to weight cues displayed in the eyes more than Western Caucasian observers possibly because the eye region provides more diagnostic information of another's true emotions. In Eastern cultures, where the control of emotion to maintain harmonious relationships

with others is valued, East Asians learned to suppress their true emotion (Friesen, 1972). Therefore, fixating more on the eyes, where the muscles are more difficult to control than the muscles around the mouth, may reveal more useful cue to identify East Asians' true emotions (Ekman, 1992).

Finally, results showed that for the face recognition task, the mean fixation duration on the mouth and eyes were significantly longer than the nose, whereas the mean fixation duration on the mouth was significantly longer than the eyes, followed by the nose for the emotion recognition task. This indicated that participants tended to direct more attention to the mouth and eyes when asked to recognize identities while the facial expressions of emotion rendered longer but fewer fixations on the mouth and shorter but more fixation on the eyes than nose. Interactions for the face recognition task revealed that participants tended to fixate longer on the eyes for Western Caucasian faces and the mouth for East Asian and African faces, which may have resulted in longer fixations on the mouth and eyes than the nose. As for the emotion recognition task, the mouth may have rendered significantly longer fixations than the eyes then nose because focusing on the most expressive region of the face may provide useful information for observers to categorize facial expressions accurately.

Analyses comparing eye tracking data for the emotion recognition and face recognition tasks revealed that participants showed a change in strategy by reducing fixations on the eyes and increasing fixations on the nose and mouth areas in the emotion recognition task compared to the face recognition task, providing further support that eye movement strategies are affected by the observers' goals and aims. Although a shift in strategy during the emotion recognition task enabled Malaysian Chinese participants to recognize most emotions for both own- and other-race faces

with equal proficiency, directing attention mainly towards the eye region (which contains diagnostic information about one's emotion in an Eastern context which encourages suppression of emotion expression) resulted in better recognition of certain own-race facial expressions, such as "happy" and "surprise. Chapter Three and Chapter Four have identified Malaysian Chinese observers' eye movement strategy and sensitivity in recognizing identities and emotions respectively; however, because of the methodological considerations regarding static stimuli (as presented in Chapter One), the following chapter aimed to investigate how Malaysian Chinese would perform in a dynamic face recognition task utilizing naturalistic stimuli that more accurately represent a typical social situation (i.e. muted videos of models introducing themselves).

Chapter Five

Dynamic Face Recognition

Abstract

Even though face recognition studies have showed high performance level in laboratory settings, accurate identification of faces does not seem to be replicable in some real life situations, such as eyewitness testimony. Researchers have proposed the use of dynamic stimuli to produce more ecologically valid results. Recent cross-cultural studies utilizing eye tracking methodologies have found that East Asian younger children and adults fixated longer on the eyes whereas Western Caucasian younger children and adults directed their attention more towards the mouth when viewing dynamic stimuli of models presented in different angles with different eye gaze and movements in the mouth. Although utilizing dynamic stimuli offers more visual information than static stimuli, posed facial and head movements do not accurately represent a typical social interaction. Therefore, the current study aimed to utilize more naturalistic stimuli by showing participants muted videos of African, East Asian, and Western Caucasian models giving a brief introduction of themselves. When recognizing all three races of faces, Malaysian Chinese participants fixated most on the nose, followed by the mouth then eyes. Focusing on the center of the face enabled participants to gain a more holistic view of models' facial and head movements, which proved to be beneficial in providing useful cues to recognize identities. Participants recognized all three races of faces equally well. The findings of this chapter revealed a fixation pattern different from those employed in a static face recognition task utilizing an identical procedure (see Chapter Three), which imply that observers' eye movement strategy is affected by the type of stimuli presented.

Hence, utilizing stimuli that more accurately represent a typical social interaction may produce results that are more valid.

5.1. Introduction

Static stimuli, like illustrated emoticons (e.g. Yuki, Maddux, & Masuda, 2007) or photographs of posed (e.g. Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Jack, Blais, Scheepers, Schyns, & Caldara, 2009) or natural faces (e.g. Matsumoto & Willingham, 2009), have been commonly used for research involving face processing. However, a natural social encounter with another individual involves dynamic facial motion and body movements that convey substantial amounts of social information and visual cues about one's external and internal state (e.g. nodding or shaking one's head in agreement or disapproval, eye gaze changes that show interest or boredom, facial expressions of emotion, etc.). Moreover, research has found that high performance levels in laboratory studies of face recognition does not seem to be replicable in some real life situations, such as eyewitness testimony (Wells & Olsen, 2003; Meissner & Brigham, 2001). Therefore, utilizing dynamic stimuli in face processing studies would allow researchers to explore related topics, such as face recognition, in a more naturalistic way, which may be a more accurate representation of a typical social situation.

5.1.1. Facial Motions and Social Signals

Facial movements can be identified as either rigid or non-rigid motions; both of which provide different views of a face for an observer to better construct a more complete representation of the face and derive social information about the person. Rigid motions, such as nodding, shaking, tilting, and rotating of the head, provide

social interaction signals to another person, whereas non-rigid motions, which can be categorized into facial expressions, eye gaze changes, and facial speech, may distort the identifying features of a face.

Even though facial movements provide ample visual information about a face by conveying rich social cues, perceiving dynamic faces may prove challenging to observers as these visual cues not only need to be constantly examined and interpreted, they may also alter the facial structure that is crucial for recognizing unfamiliar faces (O'Toole, Roark, & Abdi, 2002; O'Toole et al., 2006). Therefore, researchers aimed to examine how dynamic stimuli would affect observers' recognition accuracy and how facial motion may affect face recognition.

5.1.1.1. The Representation Enhancement and Supplemental Information Hypotheses

There are two hypotheses about the role of facial motion in face recognition. The representation enhancement hypothesis proposes that facial motions enhance the encoding of the invariant structure of a face by providing observers additional structures of a face. Hence, viewing dynamic stimuli would be beneficial for face recognition because movements would allow observers to build a three-dimensional representation of a face (O'Toole et al., 2002; O'Toole et al., 2006; Knappmeyer, Thornton, & Bulthoff, 2003).

The supplemental information hypothesis suggests that facial motions provide additional information to the invariant structure of the face by reflecting characteristics that are specific to an individual, such as one's characteristic crooked smile or a head tilt. Studies showed that these dynamic facial signatures enhanced participants' recognition of familiar or famous

faces, especially in non-optimal viewing conditions, but did not improve participants' recognition of unfamiliar or unknown faces as observers may need time to learn the dynamic facial signatures that provide reliable identification information (Knight & Johnston, 1997; Lander, Christie, & Bruce, 1999; Lander & Bruce, 2000).

Previous studies examining participants' recognition accuracy when viewing static images and dynamic stimuli of famous and unknown faces provide evidence that participants are better at recognizing moving than still faces (Knight & Johnston, 1997). More specifically, facial movements improved recognition of photographically negative faces which were presented in an upright position, but movement did not affect recognition accuracy for upright or inverted faces when they were presented in an optimal viewing condition (i.e. faces which were presented as photographic positives).

Studies have also investigated how different types of stimuli may affect observers' recognition for degraded images of famous faces (Lander & Bruce, 2000). Participants either watched a 2.5 seconds moving sequence of nine images, a reversed presentation of the moving sequence, a speeded up clip of the original forward-moving clip, or a single static image of faces. Results showed that participants were best at recognizing the original forward-moving clip presented at normal speed, whereas reversed presentation and speeding up of the clip decreased participants' recognition ability. It was also found that recognition performance was better in all three dynamic stimuli conditions than the static condition. The findings of this experiment imply that the timing (i.e. normal speed or speeded up) and direction (i.e. forward-moving or reversed) of dynamic stimuli are both crucial factors in enhancing

face recognition. Additionally, the results also supported the supplemental information hypothesis, providing evidence that facial motions provide additional information that may reflect characteristics specific to an individual, resulting in better recognition performance for dynamic faces.

In another study, participants studied a list of 36 names of famous people in 2 minutes (Lander et al., 1999). If participants were familiar with at least 27 names, they then proceeded to the testing phase, in which they were presented with either upright or inverted static and dynamic faces either photographically positive or negative. Each face appeared for 2.5 seconds and participants verbally named the faces to the experimenter. Results showed that participants were better at recognizing moving faces under all conditions, including inverted and negative faces. A follow up experiment included moving and static sequences of nine frames. The nine-frame moving sequence was repeated thrice within each trial, for a total of 3.24 seconds per trial. For the static ordered condition, all nine frames were presented on the screen at the same time, in a consecutive order of the moving sequence for 3.24 seconds. As for the static jumbled condition, the nine frames were, again, presented all at once for 3.24 seconds, but in a randomized order. All faces were upright and thresholded (i.e. images were converted to grey levels and then posterized) or inverted and photographically positive. Results revealed higher recognition accuracy for moving than static sequences, implying that the dynamic movement of a face, rather than the amount of static information, enhances observers' recognition accuracy since both groups of participants were presented with an identical amount of static information.

Studies exploring the role of head and facial movements on recognition ability discovered that rigid and non-rigid head movements provide useful cues for observers to categorize identity and gender respectively (Hill & Johnston, 2001). In this experiment, a three-dimensional average head was computer-animated with different movements that were captured from twelve real people. Rigid head movements (i.e. head rotation and translation without changing the shape of the head) were separated from non-rigid head movements (i.e. change of expressions without movement of the head) to examine how different types of motions affect the perception of faces. Results showed that categorization accuracy for identities was higher for participants who watched animations with rigid head movements than those with non-rigid head movements; possibly because non-rigid motions were uninformative in providing differentiating cues for identities as all "actors" were performing similar movements (i.e. telling jokes). However, non-rigid head movements were particularly useful for categorizing one's gender. Categorization accuracy for both identity and gender were reduced significantly when animated faces were inverted, indicating that dynamic information may be processed by a model-based system that is specialized for upright faces (Hill & Johnston, 2001).

5.1.2. Eye Tracking

Developmental studies revealed that Caucasian infants (age ranging from 184 to 300 days) fixated longest on the eyes, followed by the mouth and nose when viewing videos of three Caucasian and three African adult females holding a neutral expression and counting forward for 30 seconds. As age (in days) increased,

Caucasian infants showed a diversion of attention from the mouth to the eyes when viewing own-race faces; the fixation time for the eyes and mouth was similar across age for other-race faces but the fixation time was longer for the eyes and shorter for the mouth when older infants viewed own-race faces, while the fixation time on the nose did not differ as a result of age or race of face (Wheeler, Anzures, Quinn, Pascalis, Omrin, & Lee, 2011). Conversely, studies involving Chinese infants (age ranging from 122 to 284 days) who watched videos of five Chinese and five Caucasian adult females holding a neutral expression and counting forward for 30 seconds showed that with increased age, Chinese infants directed equal attention to the nose (across age) for own-race faces but fixation time landing on the nose was decreased for the other-race Caucasian faces, while attention to the eyes and mouth was maintained for own- and other-race faces (Liu, Quinn, Wheeler, Xiao, Ge, & Lee, 2011). Both studies involving infants found an own-race bias on face scanning, a finding contrasting with other studies involving older children and adults using either static images or dynamic stimuli which reported similar eye movement strategy when viewing own- and other-race faces (Blais et al., 2008; Jack et al., 2009; Kelly, Liu, Rodger, Miellet, Ge, & Caldara, 2011b; Kelly, Jack, Miellet, De Luca, Foreman, & Caldara, 2011a; Senju, Vernetti, Kikuchi, Akechi, & Hasegawa, 2013a; Senju, Vernetti, Kikuchi, Akechi, Hasegawa, & Johnson, 2013b)

In a study involving young children between the ages of one to seven years, British and Japanese participants viewed 6-second stimuli of two Caucasian and two East Asian models (Senju et al., 2013a). Stimuli showed different angles of the models' face (i.e. upright and turned 30 degrees to the left or right) with direct or averted gaze when posing a neutral expression or when smiling. Eye tracking results showed that British young children fixated longer on the mouth than Japanese young

children, a finding similar with previous studies using static images (Blais et al., 2008; Jack et al., 2009), while Japanese young children directed their attention more towards the eyes, a finding comparable with those reported in a facial expressions task (Jack et al., 2009).

British and Japanese adults performed similarly when viewing computer-generated avatar faces of two Caucasian and two East Asian models presented in different angles (i.e. upright and turned 30 degrees to the left or right) with direct or averted gaze and movements in the mouth (i.e. smiling, mouth opening, or closed mouth). Identical with the fixation pattern employed by younger children (Senju et al., 2013a), British participants tended to fixate more on the mouth than Japanese participants. Japanese participants, on the other hand, tended to fixate on the eyes longer than British participants, with comparable fixation duration landing on the eye closer to the observer for both group of participants, but longer fixations on the other eye for Japanese participants only (Senju et al., 2013b). However, both studies involving young children and adults did not find a preference for the nose in East Asian participants, contrasting with the fixation pattern previously reported in face recognition studies (Blais et al., 2008). Senju et al. (2013b) suggested that the nature of the task (e.g. face recognition) may influence observers to employ different strategies to analyze the face and, hence, may not occur when viewing dynamic stimuli passively. It was also suggested that East Asian observers may focus more on the center of the face when models are holding neutral expressions, but direct their attention towards the eyes when the face is expressive (Senju et al., 2013b).

5.2. Current Study

Previous eye tracking studies mostly involved passive viewing of dynamic stimuli that do not accurately represent a typical social situation (e.g. avatar faces or posed facial and head movements). Therefore, the current study aimed to use a more naturalistic stimuli (i.e. models giving a brief introduction of themselves) to examine Malaysian Chinese observers' sensitivity and eye movement strategy in recognizing muted videos of African, East Asian, and Western Caucasian identities.

If Malaysian Chinese utilize a similar fixation pattern when viewing static and dynamic stimuli, participants would focus mainly on the eye and nose region when recognizing faces. Employing a similar eye movement strategy when viewing dynamic stimuli in a face recognition task may result in similar behavioral results; with higher sensitivity in recognizing the more familiar East Asian and Western Caucasian faces than African faces. However, previous face processing studies involving dynamic stimuli in naturalistic settings have revealed a different fixation pattern, with more fixations directed towards the mouth when observers are making judgments about scenarios based on recorded videos of models' reactions (Pillai, Sheppard, & Mitchell, 2012). Therefore, it is plausible that Malaysian Chinese participants' fixation pattern may alter when viewing dynamic stimuli as facial movements may redirect participants' attention to features that may be more animated or contain more visual information. Again, participants may be better at recognizing East Asian and Western Caucasian faces, with which they are more familiar.

5.3. Methods

5.3.1. Participants

Twenty-three ethnically Chinese students attending Sunway University (11 males, 12 females, mean age 19.48 years) participated in this study. The twenty-two Malaysian participants had not lived outside of Malaysia at all; the one female Singaporean participant was born in Singapore but has lived in Malaysia for 18 years. All participants had normal or corrected vision and were given a bar of chocolate for their participation. Written informed consent was obtained from all participants and the protocol was approved by the University of Nottingham, Faculty of Science Ethics Committee.

5.3.2. Materials

Sixty "models" (term used as short-hand—the volunteers were not trained models) were video recorded while they gave brief introductions of themselves. Models were distributed equally among race (20 African, 20 East Asian, and 20 Western Caucasian with a mean age of 22.4 years, 20.1 years, and 24.7 years respectively) and gender, with a mean age of 22.4 years. A one-way analysis of variance was conducted on participants' age, revealing a significant difference of age between races, $F(2, 57)=4.29$, $p=.018$. Post-hoc tests using Bonferroni showed that Western Caucasian models were significantly older than East Asian models ($p=.015$); however, the age of African models were not significantly different from Western Caucasian ($p=.446$) and East Asian ($p=.446$) models. Models were students or staff in the University of Nottingham Malaysia Campus who participated voluntarily. Written informed consent was obtained from all models and the protocol was approved by the University of Nottingham, Faculty of Science Ethics Committee.

A Panasonic HDC-TM300 video camera was used to record the models in a Psychology laboratory room. Positioned at approximately 1.7 meters on a tripod, models' face, neck, and shoulders were included in the frame. The models sat on a chair against a white background and were asked to introduce themselves by saying their name, age, where they come from, what course they were studying or what course they were teaching or their hobby.

All videos were transferred from the video camera to a HP Compaq LE1902x desktop. The videos were muted to avoid recognition of identities based on other factors, such as contextual or auditory information, and were cut into two segments (i.e. the first five seconds of 30 videos were presented in the learning phase and the last five seconds of all 60 videos were presented in the recognition phase) using Avidemux 2.6 (32-bit). A gray mask was then applied to all videos to prevent participants from recognizing the identities based on the models' clothing using Adobe Premiere CS6 (see Figure 5.1). The videos, which were 550x440 pixels in size, were presented on a Tobii T60 eye tracker (a 17in TFT monitor with a screen resolution of 1280x1024 pixels) at a distance of approximately 60cm. Presentation of the stimuli was controlled by the Tobii Studio software, version 2.3.

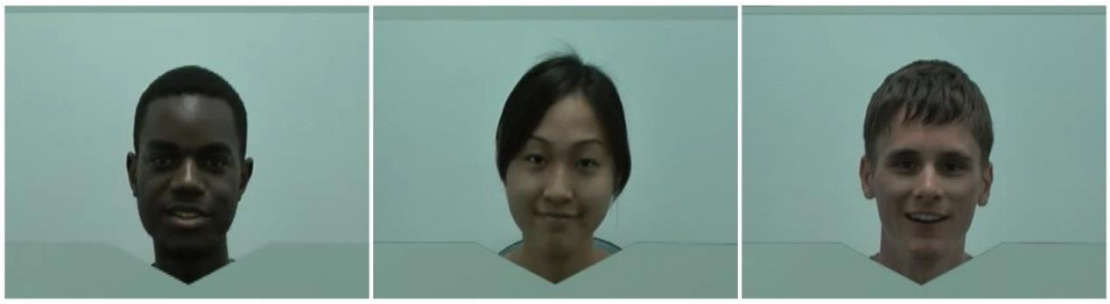


Figure 5.1. Examples of stimuli used in the experiment.

5.3.3. Procedure

The experiment involved two phases: the learning phase and the face recognition phase. During the learning phase, 30 videos (5 male, 5 female East Asian, 5 male, 5 female Western Caucasian, and 5 male, 5 female African) were shown on a Tobii eye tracker and participants were asked to rate the faces for attractiveness on a seven-point Likert scale. Participants were not informed that they would be performing a subsequent face recognition task. To distract participants from remembering the faces, participants filled out the Autism Quotient questionnaire as a distractor task. Upon completion of the questionnaire, participants performed a face recognition task in which 60 videos (10 male, 10 female East Asian, 10 male, 10 female Western Caucasian, and 10 male, 10 female African), of which half were new faces, were presented. Participants gave a yes or no response to indicate if they had seen the face before.

On each trial, a central fixation cross was presented pseudorandomly in one of four quadrants of the computer screen to avoid fixation bias for one second, followed by a video presented in the center of the screen. Each video was presented for 5 seconds in both phases and was followed by a forced-choice question in relation to the task (e.g. "How attractive do you think this face it?" for the learning task and "Have you seen this picture before?" for the recognition task). Each response was subsequently followed by a central fixation cross, which preceded the next video.

5.3.4. Eye Tracking

Eye tracking information and the calibration procedure were identical to those presented in Chapter Three.

5.3.5. Data Analyses

The procedure for processing and analyzing the behavioral and eye tracking data was identical to those presented in Chapter Three.

To ensure the validity of the eye tracking data, only participants who had an average number of fixation equal to or more than 5 counts during the recognition task were included in the data set. Six participants (5 males and 1 female) did not meet the criteria and were excluded from the eye tracking analysis.

For the Linear Mixed Modeling analyses, participant ID nested within task and face ID nested within race were included as random factors for total fixation count and mean duration. For total fixation duration, only participant ID nested within task was included as a random factor because the final Hessian matrix was not positive, which casts doubt on the validity of results; however, all results with or without face ID nested within race included as a random factor was identical. No random factors were included for total number of first fixations because of the same reason; again, all results with or without participant ID nested within task and face ID nested within race included as random factors were identical.

5.4. Results

5.4.1. Recognition Sensitivity

One-sample t-tests on A' values showed that Malaysian Chinese participants were able to recognize dynamic African, $t(22)=5.94$, $p<.001$, East Asian, $t(22)=5.11$, $p<.001$, and Western Caucasian faces, $t(22)=6.13$, $p<.001$, above chance level (test value = 0.5).

Table 5.1. Hit rates and false alarm rates in recognizing African, East Asian, and Western Caucasian faces.

	Hit Rates (SD)	False Alarm Rates (SD)
African faces	0.73 (0.18)	0.17 (0.14)
East Asian faces	0.75 (0.18)	0.12 (0.12)
Western Caucasian faces	0.79 (0.18)	0.16 (0.15)

A one-way repeated measures analysis of variance (ANOVA) was conducted on participants' recognition sensitivity (A'). Participants had higher hit rates for Western Caucasian faces, followed by East Asian then African faces (see Table 5.1); however, A' values did not differ significantly between races of faces, $F(1.52, 33.34)=.89$, $p=.395$, Greenhouse-Geisser corrected (see Figure 5.2).

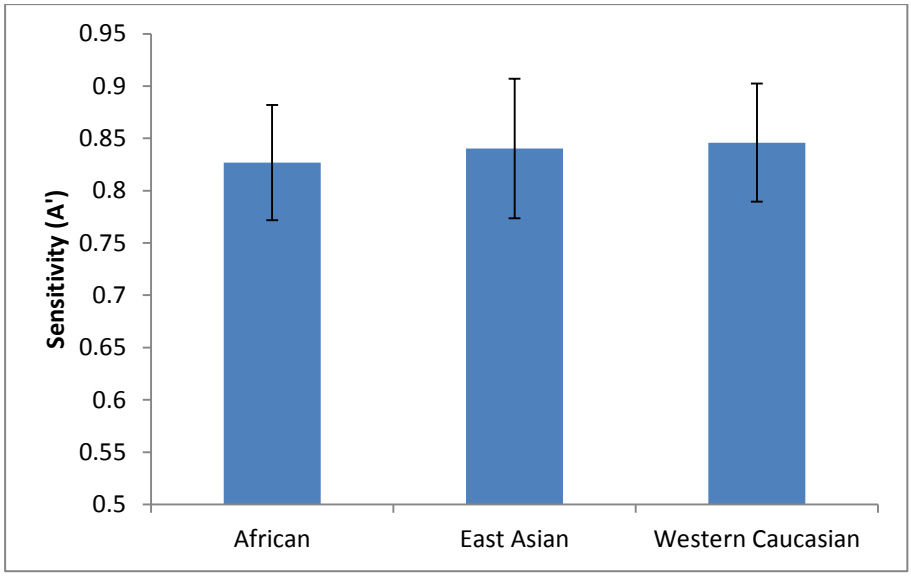


Figure 5.2. Malaysian Chinese participants' recognition sensitivity for African, East Asian, and Western Caucasian faces. Error bars report standard errors of the mean. Participants recognized all three races equally well.

5.4.2. Eye Tracking Results

Linear Mixed Modeling analyses were conducted to test the effect of task (learning and recognition), race of face (African, East Asian, or Western Caucasian) and feature of face (eyes, mouth, or nose) on the four dependent variables: total number of first fixations, total fixation count, total fixation duration, and mean fixation duration, see Table 5.2 for a summary of statistics for each dependent variable. Analyses for total fixation duration revealed a pattern similar to those observed in the total fixation count; therefore, the results for total fixation duration were not discussed in the current chapter (refer to Appendix D for the results of total fixation duration).

Table 5.2. Summary of statistics for each dependent variable in the linear mixed models.

	Total Number of First Fixations	**Total Fixation Count**	**Total Fixation Duration**	**Mean Fixation Duration**
Task	$F(1, 4572)=.35$, $p=.553$	$F(1, 32.54)=.35$, $p=.559$	$F(1, 32.07)=.39$, $p=.537$	$F(1, 32.60)=.01$, $p=.944$
Race	$F(2, 4572)=.02$, $p=.980$	$F(2, 61.60)=1.48$, $p=.235$	$F(2, 4540.06)=2.41$, $p=.090$	$F(2, 44.98)=.80$, $p=.457$
Feature	$F(2, 4572)=115.73$, $p<.001$**	$F(2, 4496.20)=396.71$, $p<.001$**	$F(2, 4540.06)=394.63$, $p<.001$**	$F(2, 3522.52)=78.95$, $p<.001$**
Task x Race	$F(2, 4572)=.004$, $p=.996$	$F(2, 1305.76)=1.11$, $p=.332$	$F(2, 4540.06)=1.24$, $p=.289$	$F(2, 1197.24)=.90$, $p=.409$
Task x Feature	$F(2, 4572)=2.23$, $p=.108$	$F(2, 4496.20)=.39$, $p=.676$	$F(2, 4540.06)=.26$, $p=.769$	$F(2, 3519.72)=.59$, $p=.557$
Race x Feature	$F(4, 4572)=2.06$, $p=.083$	$F(4, 4496.20)=.86$, $p=.487$	$F(4, 4540.06)=.47$, $p=.755$	$F(4, 3509.36)=.88$, $p=.478$
Task x Race x Feature	$F(4, 4572)=.32$, $p=.866$	$F(4, 4496.20)=.554$, $p=.696$	$F(4, 4540.06)=.39$, $p=.815$	$F(4, 3507.04)=.34$, $p=.850$

* for $p<.05$

** for $p<.001$

5.4.2.1. Total Number of First Fixations

For total number of first fixations, there was a main effect of feature, $F(2, 4572)=115.73$, $p<.001$, and a trend towards an interaction between race and feature $F(4, 4572)=2.06$, $p=.083$ (see Figure 5.3). Post-hoc tests using Bonferroni Pairwise Comparisons showed that participants first fixated on the nose significantly more than the eyes ($p<.001$) and mouth ($p<.001$), and the eyes significantly more than the mouth ($p<.001$).

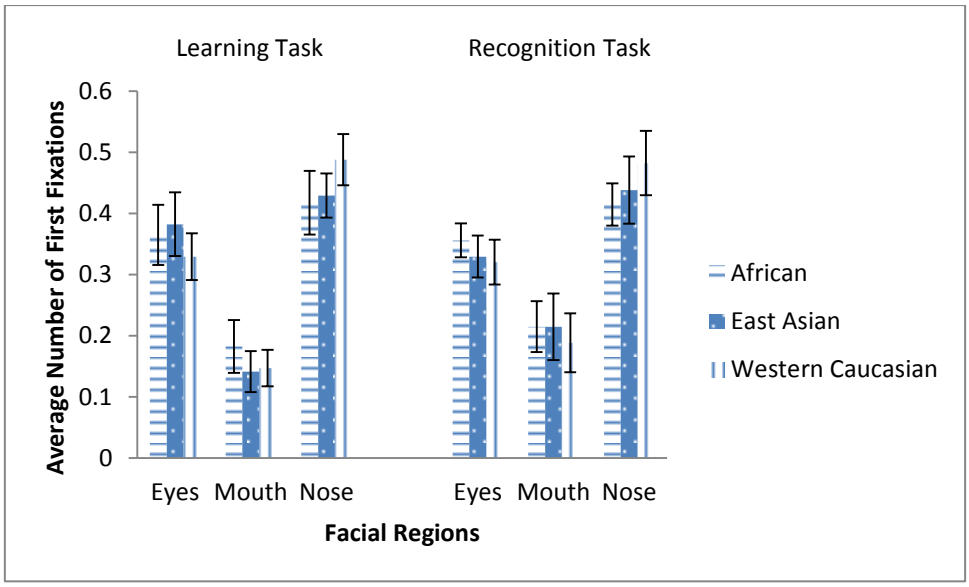


Figure 5.3. Average number of first fixations landing on the eyes, mouth, and nose during the learning and recognition tasks. Error bars report standard errors of mean. Participants tended to first fixate on the nose, followed by the eyes, then mouth.

To examine the trend towards an interaction between race and feature, separate Linear Mixed Models were conducted for each of the three races. Participant ID, face ID, and feature were included as factors (dependent variable = total number of first fixations; fixed factor = feature of face). There was a main effect of feature for African, $F(2, 1527)=28.87$, $p<.001$, East Asian, $F(2, 1527)=37.57$, $p<.001$, and Western Caucasian faces, $F(2, 1527)=59.83$,

p<.001. Post-hoc tests using Bonferroni Pairwise Comparisons showed that for East Asian and Western Caucasian faces, participants tended to direct their attention first to the nose significantly more than the eyes (p=.006 for East Asian faces, p<.001 for Western Caucasian faces) and mouth (p<.001 for both races), and the eyes more than the mouth (p<.001 for both races). For African faces, participants first fixated on the nose (p<.001) and eyes (p<.001) significantly more than the mouth; however, the total number of first fixations landing on the nose and eyes did not differ significantly (p=.147). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and race were included as factors (dependent variable = total number of first fixations; fixed factor = race of face). There was a trend towards a main effect of race for the nose only, F(2, 1527)=2.58, p=.076; there were no significant differences between the total number of first fixations landing on the eyes, F(2, 1527)=.73, p=.480, and mouth, F(2, 1527)=.72, p=.488, for all three races of faces. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that there was a trend towards more first fixations landing on the nose for Western Caucasian than African faces (p=.082); however, the total number of first fixations landing on the nose for East Asian faces did not differ from Western Caucasian (p=.346) and African faces (p=1).

5.4.2.2. Total Fixation Count

For total fixation count, there was a main effect of feature only, F(2, 4496.2)=396.714, p<.001 only (see Figure 5.4). Post-hoc tests using Bonferroni Pairwise Comparisons showed that participants fixated on the nose

significantly more than the mouth (p<.001) and eyes (p<.001), and the mouth significantly more than the eyes (p<.001).

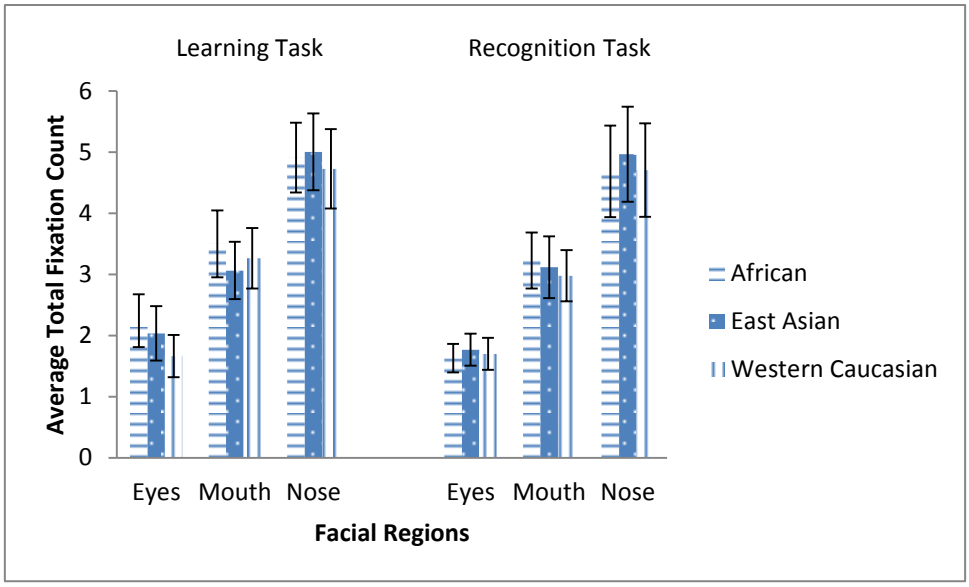


Figure 5.4. Average total fixation count for the eyes, mouth, and nose during the learning and recognition tasks. Error bars report standard errors of mean. Participants fixated on the nose significantly more than the mouth and eyes, and the mouth significantly more than the eyes.

5.4.2.3. Mean Fixation Duration

For mean fixation duration, there was a main effect of feature only, F(2, 3522.522)=78.945, p<.001 (see Figure 5.5). Post-hoc tests using Bonferroni Pairwise Comparisons showed that the mean fixation duration for the mouth was significantly longer than the nose (p=.001) and eyes (p<.001), and the nose significantly longer than the eyes (p<.001).

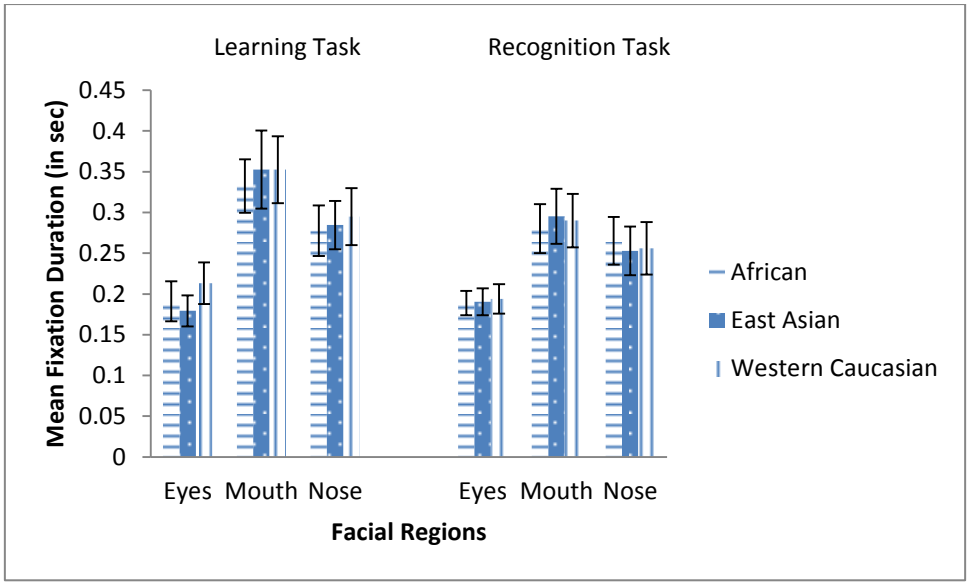


Figure 5.5. Mean fixation duration for the eyes, mouth, and nose during the learning and recognition tasks, in seconds. Error bars report standard errors of mean. The mean fixation duration for the mouth was significantly longer than the nose and eyes, and the nose significantly longer than the eyes.

5.5. Discussion

The current study investigated Malaysian Chinese participants' recognition sensitivity and eye movement strategies when viewing muted videos of African, East Asian, and Western Caucasian identities introducing themselves. The primary findings were that 1) Malaysian Chinese participants recognized African, East Asian and Western Caucasian videos equally well, 2) attention was directed first to the nose more than the eyes, followed by the mouth, 3) participants fixated most frequently on the nose, followed by the mouth then eyes, and 4) the mean fixation duration for the mouth was significantly longer than the nose, followed by the eyes.

As discussed in Chapter Three, Malaysian Chinese participants recognized the familiar East Asian and Western Caucasian faces better than the less familiar African

faces when perceiving static images. Previous studies examining the role of dynamic stimuli on face recognition have found better recognition performance when identifying famous faces under non-optimal viewing conditions (Knight & Johnston, 1997; Lander & Bruce, 2000; Lander et al., 1999); the current study aimed to investigate whether presenting a more naturalistic dynamic stimuli under optimal viewing conditions may have an impact on the other-race effect, a robust phenomenon that has been found reliable across various experimental conditions (Meissner & Brigham, 2011; Bothwell et al., 1989). Behavioral results revealed an equal ability in recognizing African, East Asian and Western Caucasian videos. The advantage of recognizing all three races of faces with an equal ability is not due to ceiling effects as the average A' values for the current dynamic face recognition task (M=0.84, SD=0.28) is similar to those for the static face recognition task (M=0.88, SD=0.08) reported in Chapter Three. Moreover, the hit rate (i.e. percentage of correct recognition responses) was less than 80% across conditions, which does not reach ceiling. This indicates that dynamic stimuli may, perhaps, reduce the other-race effect by providing more diagnostic information about a face that is useful for recognizing identities.

Eye tracking results showed that for both learning and recognition tasks, the total number of first fixations on the nose was significantly more than the eyes and mouth, and the eyes more than the mouth. This pattern is similar to those observed in the static face recognition task presented in Chapter 3, indicating that participants consistently fixate first on the center of the face to gain a more holistic view at first glance despite a change in stimuli. Additionally, participants made more first fixations on the nose, followed by the eyes then mouth for East Asian and Western Caucasian faces, while there were more first fixations on the nose and eyes than the mouth for

African faces. This indicates that at first glance, participants may rely on different features to make judgments about faces of different races.

As compared with the previous static face recognition study that utilized an identical procedure, Malaysian Chinese participants employed a different eye movement strategy to perceive static images and dynamic videos. Participants tended to focus on the eyes, followed by the nose then mouth when viewing images of faces, while participants directed their attention to the nose, followed by the mouth then eyes when viewing videos of faces. Participants may focus mainly on the eyes when perceiving images as the eyes may provide differentiating cues for recognizing faces in static stimuli (Ellis, Deregowski, & Shepherd, 1975); however, when presented with videos, participants may choose to focus on the center of the face to gain a more holistic view of facial and head movements that have been shown to provide useful cues in recognizing identities (Hill & Johnston, 2001). Additionally, there was an increase in fixations for the mouth in the dynamic face recognition task, a pattern consistent with previous eye tracking study investigating participants' judgments about scenarios based on recorded videos of models' natural reactions (Pillai et al., 2012). Although it is possible that participants may be attracted to the mouth based on visual saliency, previous eye tracking study has suggested that fixations are modulated by the degree of task-relevant information a feature provides, indicating that the mouth may provide useful cues for identifying faces (Vo, Smith, Mital, & Henderson, 2012). It is unlikely that participants were trying to lip read the muted videos as a previous study has found that participants tended to fixate less on the mouth when viewing muted videos, as compared to videos with audio, when judging the likeability of naturalistic videos (Vo et al., 2012). Moreover, lip reading would not help participants succeed in the current task as different segments of the videos were

presented in the learning and recognition phase. These findings suggest that although the eyes may provide useful cues in recognizing static images of faces, focusing on the lower regions of the face (i.e. the nose and mouth area) would allow observers to gain a more holistic view of the models' facial and head movements and may provide important cues that are advantageous for recognizing dynamic faces.

The mean fixation duration revealed that participants made longer fixations on the mouth than the nose and eyes, and the nose longer than the eyes, implying that participants made fewer but longer fixations on the mouth. The mean fixation duration identified in the current dynamic face recognition task differed from those reported in the static face recognition task; more specifically, participants tended to make shorter fixations on the eyes when viewing dynamic faces. Although participants tended to fixate more frequently on the center of the face to gain a more holistic view of models' facial and head movements, participants fixated longer on the mouth to process cues advantageous for recognizing dynamic faces. Participants' tendency to focus on the mouth is similar with the pattern observed in previous eye tracking study involving naturalistic dynamic stimuli, affirming that the focus on the mouth area may be ideal for processing faces.

Previous eye tracking studies have found that East Asian younger children and adults tended to fixate longer on the eyes whereas Western Caucasian younger children and adults directed their attention more towards the mouth when viewing dynamic stimuli of models presented in different angles with different eye gaze and movements in the mouth (Senju et al., 2013a; Senju et al., 2013b). In comparison, Malaysian Chinese participants showed a different fixation pattern that does not seem well-matched with either the Eastern or Western pattern, with longer total fixation duration on the nose, followed by the mouth then eyes. However, the current eye

tracking results are not directly comparable with previous eye tracking studies because of the different nature of the task (i.e. face recognition versus passive viewing) and stimuli (i.e. models talking naturally about themselves versus posing between neutral and happy expressions). Therefore, the findings may either reflect a unique fixation pattern that Malaysian Chinese participants employ to better recognizing own- and other-race faces or that the differences in strategies may be caused by the dissimilarities in methodology.

The findings of this chapter showed that Malaysian Chinese participants employed a different eye movement strategy when a different type of stimuli (i.e. dynamic videos instead of static images) was presented in an identical face recognition task. Participants changed their strategy according to the type of stimuli presented. The next chapter aimed to identify whether the change of stimuli would also affect the eye movement strategy that Malaysian Chinese participants employ when viewing dynamic stimuli of models describing an emotional experience ("happy," "surprise," "fear," "disgust," "anger," "sad," plus "neutral;" Ekman, 1970) and how that may affect participants' categorization sensitivity.

Chapter Six

Dynamic Emotion Recognition

Abstract

Dynamic emotion recognition studies yielded mixed results on participants' behavioral performance, possibly due to the varying methods used to develop the dynamic stimuli. A majority of these studies morphed between neutral and full-blown expressions; however, the dynamic aspects of facial expression production are not linear and the dynamic motion of morphed expressions is not the same as movements involved in real expressions. Therefore, the current study aimed to utilize more naturalistic stimuli by showing Malaysian Chinese participants muted videos of African, East Asian, and Western Caucasian models describing one of seven emotional experiences. Results showed that participants were best at recognizing "happy" while there was a trend towards "fear" being recognized worse than "anger" and "disgust." Participants fixated on the mouth and nose more than the eyes to recognize emotions, possibly because the mouth provides important cues for identifying emotions while focusing on the nose is an optimal strategy to observe the models' facial and head movements from a holistic view. Additionally, participants altered their fixation pattern when instructed to perform different tasks on an identical set of stimuli; fixations on the mouth increased in the dynamic emotion recognition task, as compared to the dynamic face recognition task, suggesting that the mouth provides more diagnostic information to recognize emotions whereas focusing on the center of the face may be a more optimal strategy to recognize identities.

6.1. Introduction

Studies investigating observers' ability to categorize facial expressions typically involve static images that capture a single portrayal of an intense emotion; however, static stimuli may not be ecologically valid as natural expressions unfold dynamically in a real world setting. Previous studies utilizing static stimuli have yielded high accuracy rates when categorizing facial expressions of emotion (Ekman & Friesen, 1971; Biehl, Matsumoto, Ekman, Hearn, Heider, Kudoh, & Ton, 1997; Jack, Blais, Scheepers, Schyns, & Caldara, 2009); however, higher accuracy rates for categorizing images of facial expressions does not necessarily reflect better recognition ability in real life situation as static stimuli lack dynamic information that is crucial in the nonverbal communication of one's emotion. Observing one's expressions of emotion during a typical social interaction convey more complex visual information than a single image that depicts one's strongest expression; thus, studies involving dynamic stimuli may unveil results better in line with everyday performance than those observed in previous studies.

6.1.1. Dynamic Advantage Hypothesis

Previous studies have found an advantage for dynamic stimuli in various face processing tasks, like gender and identity recognition (Lander & Bruce, 2000; Hill & Johnston, 2001); some studies found that this advantage is present only when faces were presented in non-optimal viewing conditions, such as blurred, negative, or inverted faces (Knight & Johnston, 1997; Lander, Christie, & Bruce, 1999). However, studies investigating participants' ability to categorize facial expressions of emotion yielded mixed results possibly due to the varying methods used to develop the dynamic stimuli (Fiorentini & Viviani, 2011; Hoffmann, Traue, Limbrecht-Ecklundt,

Walter, & Kessler, 2013). Some studies support the dynamic advantage hypothesis, with results showing higher accuracy rate in recognizing dynamic facial expressions than those displayed in static images (Wehrle, Kaiser, Schmidt, & Scherer, 2000; Ambadar, Schooler, & Cohn, 2005), while others did not find a clear advantage for dynamic stimuli (Fiorentini & Viviani, 2011; Hoffmann et al., 2013; Horstmann & Ansorge, 2009).

6.1.1.1. Supporting Evidence

A study involving synthetic images and animations (i.e. three-dimensional facial expressions displaying only lines and shapes without skin features that were generated by the Facial Action Composing Environment, which is a tool to create three-dimensional animated facial expressions) first examined the validity of the synthetic stimuli before exploring the effect of stimuli type on recognition accuracy (Wehrle et al., 2000). The first study found that participants' recognition accuracy for both static and dynamic synthetic stimuli was above chance level but yielded lower recognition accuracy than posed photographs of real people, suggesting that the synthetic stimuli appropriately included at least some of the relevant features and facial muscle movements that are crucial in emotion recognition. In the second study, participants were presented with one of three types of stimuli; the sequence condition (i.e. changes in facial expressions based on intensity of emotion or manipulating cues, such as facial and head movements, direction of gaze, or physiognomy of the face, sequentially), movement condition (i.e. simultaneous animation of all facial and head movements that result in the final expression of a target emotion), or static condition (i.e. a single image of

the final expression of a target emotion). Results showed that participants were able to recognize emotions significantly better in both dynamic conditions than the static condition; however, performance in the sequence and movement condition did not differ. This finding implies an advantage for dynamic stimuli in emotion recognition.

Another study aimed to examine whether the amount or type of visual information affects emotion recognition (Ambadar et al., 2005). When participants were presented with dynamic faces (i.e. a subtle progression from neutral to an expressive emotion), first-last images (i.e. the first and last images of an expression), single-static image (i.e. a single image of the final expression), or multi-static images (i.e. an identical series of images as those presented in the dynamic condition, but with a noise mask in between images to eliminate motion), participants were able to identify the emotions for the dynamic and first-last condition better than the single-static and multi-static conditions. Moreover, participants' confidence ratings with their judgments were higher for dynamic than first-last and multi-static stimuli, and higher for first-last than multi-static stimuli. Participants' lower performance in the multi-static condition suggests that the advantage of dynamic stimuli may not be a result of the amount of visual information available, but rather a perception of change as evident in the shift between the first and last images.

6.1.1.2. Conflicting Evidence

On the other hand, some studies did not find a clear advantage for dynamic stimuli (Fiorentini & Viviani, 2011; Hoffmann et al., 2013; Horstmann & Ansorge; 2009). In a study by Fiorentini and Viviani (2011),

expression blends were created by morphing pairs of posed expressions from the same actor and were presented in images or high-speed video clips. Results revealed that participants were more accurate in identifying emotions for the static than dynamic condition. Results also showed that response times were shorter for the static than dynamic stimuli. Because all but one response for the dynamic stimuli were given before expressions were most intense, the authors applied a threshold model to estimate the effectiveness of integrating discriminative information by using the time course of the intensity of each action unit according to the Facial Action Coding System. Results revealed that static stimuli were processed more effectively than dynamic stimuli, suggesting no advantage for categorizing facial expressions of emotion while viewing dynamic stimuli. However, this could also imply that the dynamic aspects of facial expression production are not linear; a morph between neutral and expressions is not the same motion as the movement involved in real expressions.

When participants were presented with static and dynamic stimuli, results showed that they were able to recognize "fear" and "surprise" faces better in dynamic sequences; "happy," on the other hand, was recognized less well in the dynamic condition (Hoffmann et al., 2013). When the transformation of facial expressions (e.g. neutral to full-blown emotion) was only shown in either the upper or lower part of the face, results showed that focusing on the relevant part of the face was sufficient for participants to correctly recognize different facial expressions of emotion; more specifically, "happy," "disgust," and "sad" faces were better recognized when participants were presented with expression changes in the lower part of the face, whereas

"fear" and "surprise" were better recognized when changes were shown in the upper part of the face. Presentation of expression changes in the upper or lower part of the face did not affect participants' recognition accuracy for "angry" faces.

Search efficiency is higher for processing negative or threatening stimuli than positive or beneficial stimuli possibly because of the social and evolutionary advantages they hold (Volstorf, Rieskamp, & Stevens, 2001; Mealey, Daood, & Krage, 1996). In a series of studies by Horstmann and Ansorge (2009), participants were shown dynamic and static Musterle faces (i.e. computer generated synthetic faces displaying natural facial expressions) or illustrated emoticons portraying either positive (i.e. friendly) or negative (i.e. angry) emotions. Participants were asked to identify the presence or absence of either positive target faces among negative distractor faces or negative target faces among positive distractor faces by pressing on a corresponding key. Results showed higher search efficiency for dynamic than static Musterle faces; furthermore, the advantage for a negative target face among positive distractor faces was greater for dynamic than static stimuli. The authors speculated that stronger movement (i.e. vertical extension) in negative than positive faces may have contributed to participants' efficient search in dynamic negative faces. When movement was controlled for in illustrated emoticons and Musterle faces, results showed no advantage in search efficiency for dynamic stimuli. The advantage for a dynamic negative target face among dynamic positive distractor faces also diminished when the movement in negative and positive faces was equated.

The employment of varying methods to develop dynamic stimuli produced mixed results in participants' ability to categorize facial expressions of emotion. Therefore, it is crucial for researchers to standardize the development of dynamic facial expression stimuli which closely represent a real life situation to produce ecologically valid results.

6.1.2. Differences in the Portrayal and Processing of Facial Expressions

Research has revealed differences in the way observers portray and process different types of facial expressions. For instance, studies have identified observable differences in the portrayal of posed and natural expressions (Hess & Kleck, 1990) while brain-imaging studies examining the processing of facial expressions have found activations in different regions of the brain when viewing static and dynamic facial expressions (Kessler, Doyen-Waldecker, Hofer, Hoffmann, Traue, & Abler, 2011); both of which suggest that previous studies utilizing dynamic stimuli of posed expressions may not be ecologically valid.

In a study by Hess and Kleck (1990), participants' spontaneous expressions were video recorded while watching surgery and comedy videos under the presumption that they were unobserved. Participants then watched the same videos again and were instructed to deceive others about their emotional state while being video recorded (i.e. to display an incongruent facial expression to the actual emotion elicited by the videos watched). Results showed that the speed of onset and offset was slower for spontaneous than deliberate expressions and the degree of irregularity (e.g. pauses and changes in intensity) varied more in deliberate than spontaneous expressions (Hess & Kleck, 1990). These results indicate that there are observable differences in posed and natural facial expressions of emotion; hence, eliciting natural

expressions that closely represent a typical social interaction may yield more ecologically valid results, which may differ from previous studies involving posed expressions.

Brain-imaging studies using fMRI found that different regions of the brain are activated when participants process static and dynamic facial expressions of "happy," "fear," "disgust," and "sad" (Kessler et al., 2011). Results showed that when participants were viewing static stimuli, fMRI activation increased in the primary visual cortex, while the visual area V5 (which processes biological motion), parts of the superior temporal gyrus in the area of the superior temporal sulcus (also responsible for processing biological motion), the fusiform gyrus (also known as the "fusiform face area" was speculated to play a specific role in processing faces or objects of individual expertise), and other network of brain regions (i.e. dorsolateral prefrontal cortex, middle frontal gyrus, precentral gyrus, and the inferior parietal lobule) had higher activation when viewing dynamic stimuli. The authors speculated that dynamic faces may contain more interesting and biologically relevant information that require more visual attention; hence, resulting in a higher activation of multiple regions of the brain. These results suggest that the findings from studies utilizing static stimuli may not be an accurate representation of real life interactions as the brain processes both types of stimuli differently.

A recent study aimed to identify participants' eye movement strategies when performing two different tasks (i.e. the speech and emotion tasks) in the noise and no-noise conditions (Buchan, Pare, & Munhall, 2007). Participants were shown videos of an experienced actor saying 27 selected sentences in three emotions (i.e. "neutral," "happy," and "angry") without exaggeration. Participants were required to either report all the words they heard (i.e. the speech task) or to judge which emotion the

actor was portraying and rate the actor's portrayal of emotion on a scale of 1-9 (i.e. the emotion task) in the presence (i.e. noise condition) or absence (i.e. no-noise condition) of a commercial, multi-talker noise signal. Results showed that participants performed well in both speech and emotion tasks. In the speech task, participants performed better in the no-noise condition than in the noise condition; however, participants performed equally well in judging all three emotions in both noise and no-noise conditions in the emotion task. Eye tracking results revealed that participants varied their fixation patterns by tasks, but not by emotions; more specifically, there were more fixations on the eyes in the emotion task, as compared to the speech task, while there was a trend towards more fixations on the mouth in the speech task than in the emotion task. Participants' high accuracy in judging the emotions and speech were not found to be correlated with their fixations on the eyes and mouth respectively, suggesting that attention to the eyes and mouth may not be directly related to the successful processing of emotional and speech information. Additionally, when eye tracking data was collapsed across tasks, participants fixated more on the nose in the noise condition than the no-noise condition. These findings imply that participants adapt their eye movement strategies to maximize the amount of visual information processed according to the distribution of information on the face.

6.2. Current Study

Most studies examining observers' categorization accuracy for dynamic facial expressions of emotion lack ecological validity as the dynamic stimuli typically display a linear unfolding of a neutral face into a full-blown posed expression, which do not represent natural biological expressions encountered in real life interactions.

Varying display rules, for instance, may alter one's expression of the actual emotion felt (e.g. a posed expression of fear may be intense or even exaggerated, but one may de-intensify or mask fear with another emotion in a real-world situations; Ekman, 1970). Therefore, participants' recognition accuracy in studies employing these methods may not be an accurate representation of participants' ability to recognize emotions expressed in a typical social interaction. Moreover, previous studies have found differences in the way facial expressions are processed (Kessler et al., 2011) and portrayed (Hess & Kleck, 1990), providing more evidence for the need to develop more naturalistic dynamic stimuli that closely represent real life interactions.

To investigate Malaysian Chinese participants' sensitivity and eye movement strategy in categorizing emotions in a natural setting, the current study showed muted videos of African, East Asian, and Western Caucasian models describing one of seven emotional experiences (i.e. "happy," "surprise," "fear," "disgust," "anger," "sad," plus "neutral"). Previous studies have shown that dynamic stimuli may provide more visual information that would improve participants' categorization accuracy for facial expressions of emotion (Wehrle et al., 2000; Ambadar et al., 2005); however, no previous studies have examined participants' eye movement strategy in categorizing dynamic emotions based on models' narration. Presenting dynamic stimuli in a more naturalistic setting may be a more complex task as models may de-intensify or mask their actual emotions (e.g. laughing when talking about fearful experience) while narrating an emotional experience. Therefore, Malaysian Chinese participants may perform less well in recognizing these emotions.

Study has shown that the mouth provides the most important cue for recognizing both static and dynamic facial expressions (Blais, Roy, Fiset, Arguin, &

Gosselin, 2012). Additionally, because previous eye tracking studies on face recognition and emotion recognition tasks utilizing static stimuli revealed that Malaysian Chinese participants adopted different strategies by looking less at the eyes and more at the nose and mouth in the emotion recognition task, compared to the face recognition task (as reported in Chapter Four), Malaysian Chinese participants may also employ different fixation patterns from those reported in a dynamic face recognition task (with more fixations landing on the nose, followed by the mouth then eyes; refer to Chapter Five). Rather than fixating more on the nose, participants may direct their attention mainly towards the mouth, followed by the nose then eyes when performing a dynamic emotion recognition task as the mouth has been found to be the most diagnostic feature for identifying dynamic facial expressions (Blais et al., 2012).

6.3. Methods

6.3.1. Participants

Twenty-six Malaysian Chinese students attending Sunway University (13 males, 13 females, mean age 19.23 years) participated in this study. All participants have not lived outside of Malaysia at all. All participants had normal or corrected vision and were given a bar of chocolate for their participation. Written informed consent was obtained from all participants and the protocol was approved by the University of Nottingham, Faculty of Science Ethics Committee.

6.3.2. Materials

Sixty-five "models" (term used as short-hand—the volunteers were not trained models) were video recorded while they described two of six emotional experiences (i.e. "happy," "surprise," "fear," "disgust," "anger," and "sad"). Models were students

or staff in the University of Nottingham Malaysia Campus who participated voluntarily. Written informed consent was obtained from all models and the protocol was approved by the University of Nottingham, Faculty of Science Ethics Committee.

A Panasonic HDC-TM300 video camera was used to record the models in a Psychology laboratory room. Positioned at approximately 1.7 meters on a tripod, models' face, neck, and shoulders were included in the frame. The models sat on a chair against a white background and were randomly assigned to describe two separate experiences in which they felt two of six emotions (i.e. "happy," "surprise," "fear," "disgust," "anger," and "sad"). Models were instructed to "describe a time when you felt (the assigned emotion)" The models then rated the intensity of the emotion felt during the narrated experience and the extent to which they disclosed information about the experience on a seven-point Likert scale (i.e. 1 for "very mild" to 7 for "very intense" for the intensity of emotion and 1 for "very vague" to 7 for "very detailed" for the descriptiveness of narration), see Table 6.1 for the 42 chosen models' average self-ratings for both measures.

Table 6.1. The 42 chosen African (A), East Asian (EA), and Western Caucasian (WC) models' average self-rating of the intensity of emotion felt during the narrated experience and the descriptiveness of the narration.

	Models' Intensity Rating			**Models' Descriptiveness Rating**		
	A	**EA**	**WC**	**A**	**EA**	**WC**
Happy	4	4	3.5	4	5	4.5
Surprise	5.5	3.5	5.5	5	4	4.5
Fear	5.5	5.5	3.5	6	5.5	4.5
Disgust	5	5.5	4.5	5	5	5.5

Anger	5	3	3.5	4.5	5	5
Sad	5.5	2	2.5	3.5	2	3.5

All 130 videos were transferred from the video camera to a HP Compaq LE1902x desktop. Two raters listened to all videos without seeing the models' expressions and identified a five-second time frame in which the models sounded the most emotional and both raters rated the intensity of the emotion expressed on a seven-point Likert scale (i.e. 1 for "very mild" to 7 for "very intense"). Videos were excluded if models mentioned the emotion word (e.g. I was so "happy") in the chosen time frames. If there was a disagreement on the time frame chosen, both raters listened to the video again and discussed which segment most appropriately reflect the emotion described before coming to a mutual agreement.

A total of 42 videos were chosen; i.e. one for each emotion (i.e. "happy," "surprise," "fear," "disgust," "anger," "sad," plus "neutral" videos of models introducing themselves selected from a series of videos collected in Chapter Five), each race (i.e. African, East Asian, and Western Caucasian), and gender based on the highest average ratings given by both raters, see Table 6.2 for raters' average ratings for the 42 chosen videos. Of the selected videos, models were 7 African males and 7 African females with a mean age of 23 years, 7 East Asian males and 7 East Asian females with a mean age of 20.71 years, and 7 Western Caucasian males and 7 Western Caucasian females with a mean age of 26.5 years. A one-way analysis of variance was conducted on participants' age, revealing a significant difference of age between races, $F(2, 39)=4.02$, $p=.026$. Post-hoc tests using Bonferroni showed that Western Caucasian models were significantly older than East Asian models ($p=.023$); however, the age of African models were not significantly different from Western

Caucasian (p=.29) and East Asian (p=.819) models. Additionally, two separate 3 (Race of Face: African, East Asian, or Western Caucasian) x 7 (Emotions: "neutral," "happy," "surprise," "fear," "disgust," "anger," and "sad") one-way repeated measures analysis of variance (ANOVA) were conducted on models' self-rated intensity of emotion and descriptiveness of narration for the chosen 42 videos. No significant main effects or interactions were found for both measures, indicating that for all emotions, African, East Asian, and Western Caucasian models rated the intensity of the emotion felt during the narrated experience and the extent to which they described the experience in a similar manner.

Table 6.2. Raters' average rating of the 42 chosen African (A), East Asian (EA), and Western Caucasian (WC) models' intensity of emotion when narrating an emotional experience.

	Raters' Intensity Rating		
	A	**EA**	**WC**
Happy	6	4.75	5.5
Surprise	5.75	5	5.75
Fear	5.25	5.5	5.75
Disgust	6	5.75	6
Anger	5.75	5.75	5.25
Sad	6	5.25	6.5

All chosen videos were muted and cut into the five-second segments specified by both raters using Avidemux 2.6 (32-bit), see Figure 6.1. The videos, which were 550x550 pixels in size, were presented on a Tobii T60 eye tracker (a 17in TFT

monitor with a screen resolution of 1280x1024 pixels) at a distance of approximately 60cm. Presentation of the stimuli was controlled by the Tobii Studio software, version 2.3.

Prior to the main experiment, 5 participants (3 males, 2 females, mean age 29 years) of different racial background were asked to lip read from 7 stimuli (i.e. one of each emotion) to ensure that participants in the main experiment will not be able to recognize the emotions by lip reading. None of the participants were able to lip read a full sentence from the selected videos.



Figure 6.1. Examples of stimuli used in the experiment. Models describing the six universal emotions: “happy,” “surprise,” “fear,” “disgust,” “anger,” and “sad.”

6.3.3. Procedure

The main experiment was carried out on a Tobii eye tracker. On each of the 42 trials, a central fixation cross was presented pseudorandomly in one of four quadrants

of the computer screen to avoid fixation biases for one second, followed by a video presented in the middle of the screen. The video of an individual describing one of the seven emotions (i.e. “happy,” “surprise,” “fear,” “disgust,” “anger,” “sad,” plus “neutral”) was presented for 5 seconds, after which participants were required to identify the most appropriate emotion via a seven-alternative forced-choice procedure by pressing a corresponding response key on the keyboard. Each response was followed by the central fixation cross, which preceded the next face stimulus. At the end of the experiment, participants were asked to fill out the Autism Quotient questionnaire, which is a 50 item forced-choice self-report questionnaire that assesses autistic traits.

6.3.4. Eye Tracking

Eye tracking information and the calibration procedure were identical to those presented in Chapter Three.

6.3.5. Data Analyses

The procedure for processing and analyzing the behavioral and eye tracking data was identical to those presented in Chapter Four.

Each participant had to complete 42 trials for 5 seconds each. To ensure the validity of the eye tracking data, only participants who had an average number of fixation equal to or more than 5 counts were included in the data set. Six participants (4 males and 2 females) did not meet the criteria and were excluded from the eye tracking analysis.

For the Linear Mixed Modeling analyses, face ID nested within race was included as a random factor for total fixation count and mean fixation duration. No random factors were included for total number of first fixations and total fixation

duration because the final Hessian matrix was not positive, which casts doubt on the validity of results; however, all results with or without face ID nested within race included as a random factor were identical.

To compare how participants differed in their looking patterns when performing different tasks, eye tracking data for the current experiment was also analyzed alongside the data from the dynamic face recognition experiment presented in Chapter Five, which explored the eye movement strategies that participants used when recognizing dynamic face stimuli.

6.4. Results

6.4.1. Categorization Sensitivity

One-sample t-tests on A' values showed that Malaysian Chinese participants were able to recognize all African ($t(25)=5.38$, $p<.001$ for "neutral," $t(25)=5.38$, $p<.001$ for "happy," $t(25)=6.49$, $p<.001$ for "surprise," $t(25)=3.77$, $p=.001$ for "fear," $t(25)=8.97$, $p<.001$ for "disgust," $t(25)=8.8$, $p<.001$ for "anger," and $t(25)=5.11$, $p<.001$ for "sad"), East Asian ($t(25)=7.24$, $p<.001$ for "neutral," $t(25)=23.09$, $p<.001$ for "happy," $t(25)=4.66$, $p<.001$ for "surprise," $t(25)=9.97$, $p<.001$ for "fear," $t(25)=5.36$, $p<.001$ for "disgust," $t(25)=17.32$, $p<.001$ for "anger," and $t(25)=4.86$, $p<.001$ for "sad"), and Western Caucasian emotions ($t(25)=3.75$, $p=.001$ for "neutral," $t(25)=18.76$, $p<.001$ for "happy," $t(25)=5.59$, $p<.001$ for "surprise," $t(25)=2.34$, $p<.001$ for "fear," $t(25)=6.81$, $p=.027$ for "disgust," $t(25)=3.89$, $p=.001$ for "anger," and $t(25)=5.69$, $p<.001$ for "sad") above chance level (test value = 0.5).

A one-way repeated measures analysis of variance (ANOVA) was conducted on participants' categorization sensitivity (A') of emotions described, see Table 6.3

for participants' hit rates and false alarm rates. There was a main effect of emotion, $F(3.74, 93.52)=7.69$, $p<.001$, Greenhouse-Geisser corrected, and interaction between race of face and emotion, $F(12, 300)=3.85$, $p<.001$. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that participants were better at categorizing "happy" than "surprise" ($p<.001$), "fear" ($p<.001$), "anger," ($p=.004$), "sad" ($p<.001$), and a trend towards better categorization of "happy" than for "neutral" ($p=.064$). Additionally, there was a trend towards "fear" categorized worse than "anger" ($p=.059$), and "disgust" ($p=.071$).

Table 6.3. Hit (H) rates and false alarm (FA) rates in recognizing African (A), East Asian (EA), and Western Caucasian (WC) expressions of "neutral," "happy," "surprise," "fear," "disgust," "anger," and "sad" faces.

	H Rates (SD) for A faces	FA Rates (SD) for A faces	H Rates (SD) for EA faces	FA Rates (SD) for EA faces	H Rates (SD) for WC faces	FA Rates (SD) for WC faces
Neutral	0.58 (0.24)	0.27 (0.08)	0.50 (0.23)	0.22 (0.14)	0.62 (0.25)	0.25 (0.14)
Happy	0.58 (0.24)	0.27 (0.08)	0.65 (0.22)	0.18 (0.09)	0.67 (0.19)	0.09 (0.06)
Surprise	0.26 (0.15)	0.10 (0.07)	0.30 (0.25)	0.15 (0.12)	0.37 (0.17)	0.14 (0.08)
Fear	0.28 (0.19)	0.15 (0.09)	0.26 (0.15)	0.11 (0.08)	0.23 (0.13)	0.14 (0.09)
Disgust	0.38 (0.25)	0.08 (0.07)	0.32 (0.17)	0.12 (0.07)	0.32 (0.22)	0.10 (0.06)
Anger	0.19 (0.09)	0.08 (0.05)	0.45 (0.18)	0.07 (0.07)	0.26 (0.18)	0.12 (0.08)
Sad	0.44 (0.23)	0.18 (0.09)	0.22 (0.12)	0.13 (0.07)	0.31 (0.17)	0.13 (0.09)

In addition to the main effects, there was also an interaction between race of face and emotion. One-way within-subjects ANOVAs revealed a main effect of race for "happy," $F(2, 75)=11.50$, $p<.001$ and "anger," $F(2, 75)=14.01$, $p<.001$ emotions, while the remaining emotions were categorized equally well across races, $F(2, 75)=.07$, $p=.937$ for "neutral," $F(2, 75)=.07$, $p=.933$ for "surprise," $F(2, 75)=2.17$, $p=.121$ for "fear," $F(2, 75)=1.89$, $p=.158$ for "disgust," and $F(2, 75)=1.49$, $p=.232$ for "sad" emotions, see Figure 6.2. Post-hoc tests using Bonferroni Pairwise Comparisons showed that for "happy," participants categorized Western Caucasian ($p<.001$) and East Asian ($p=.002$) emotions significantly better than African emotions; however, the categorization sensitivity for Western Caucasian and East Asian emotions were not significantly different ($p=.902$). For "anger", participants categorized East Asian emotions significantly better than African ($p=.001$) and Western Caucasian ($p<.001$) emotions; however, the categorization sensitivity for African and Western Caucasian emotions did not differ significantly ($p=.657$). As for race, one-way within-subjects ANOVA revealed a main effect of emotion for East Asian, $F(6, 175)=8.77$, $p<.001$, and Western Caucasian faces, $F(6, 175)=5.75$, $p<.001$, but not for African faces, $F(6, 175)=1.51$, $p=.177$. Post-hoc tests using Bonferroni Pairwise Comparisons showed that for East Asian emotions, participants categorized "happy" significantly better than "neutral" ($p=.049$), "surprise" ($p=.002$), "fear" ($p=.002$), "disgust" ($p=.001$), and "sad" ($p<.001$) emotions, "anger" significantly better than "surprise" ($p=.01$), "fear" ($p=.009$), "disgust" ($p=.008$), and "sad" ($p<.001$) emotions. For Western Caucasian emotions, participants categorize "happy" significantly better than all emotions except "neutral" ($p<.001$ for "fear" and "anger," $p=.005$ for "sad," $p=.018$ for "surprise," and $p=.03$ for "disgust").

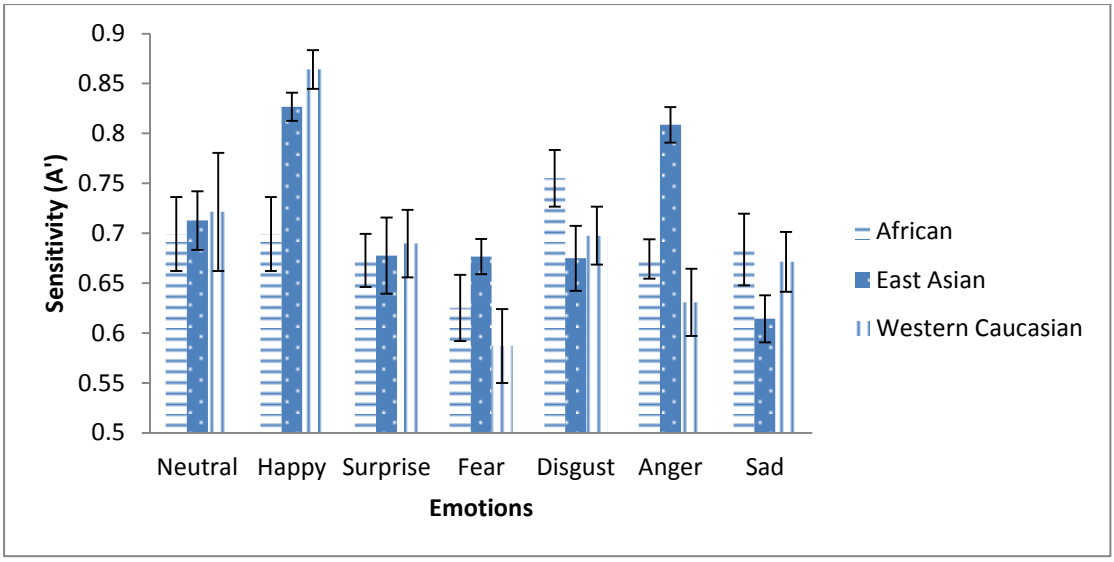


Figure 6.2. Malaysian Chinese participants' categorization sensitivity for African, East Asian, and Western Caucasian emotions. Error bars report standard errors of the mean. "Happy" was categorized better than all emotions except "neutral" and "disgust," while there was a trend towards "fear" categorized worse than "anger" and "disgust."

6.4.2. Eye Tracking Results

Linear Mixed Modeling analyses were conducted to test the effect of facial expressions of emotion ("neutral," "happy," "surprise," "fear," "disgust," "anger," and "sad"), race of face (East Asian or Western Caucasian) and feature of face (eyes, mouth, or nose) on the four dependent variables: total number of first fixations, total fixation count, total fixation duration, and mean fixation duration, see Table 6.4 for a summary of statistics for each dependent variable. Analyses for total fixation duration and mean fixation duration revealed a pattern similar to those observed in the total fixation count; therefore, the results for total fixation duration and mean fixation duration were not discussed in the current chapter (refer to Appendix E for the results of total fixation duration and mean fixation duration).

Table 6.4. Summary of statistics for each dependent variable in the linear mixed models.

	Total Number of First Fixations	**Total Fixation Count**	**Total Fixation Duration**	**Mean Fixation Duration**
Emotion	$F(6, 2457)=.03$, $p=1$	$F(6, 21)=.43$, $p=.852$	$F(6, 2457)=.214$ $p=.972$	$F(6, 21.07)=.04$, $p=1$
Race	$F(2, 2457)=.08$, $p=.925$	$F(2, 21)=2.88$, $p=.079$	$F(2, 2457)=1.17$, $p=.311$	$F(2, 21.07)=.41$, $p=.670$
Feature	$F(2, 2457)=10.42$, $p<.001$**	$F(2, 2436)=133.34$, $p<.001$**	$F(2, 2457)=134.76$, $p<.001$**	$F(2, 1719.28)=15.27$, $p<.001$**
Emotion x Race	$F(12, 2457)=.05$, $p=1$	$F(12, 21)=1.18$, $p=.355$	$F(12, 2457)=1.05$, $p=.400$	$F(12, 21.05)=.30$, $p=.982$
Emotion x Feature	$F(12, 2457)=1.38$, $p=.169$	$F(12, 2436)=.53$, $p=.895$	$F(12, 2457)=.48$, $p=.929$	$F(12, 1719.13)=.61$, $p=.835$
Race x Feature	$F(4, 2457)=.99$, $p=.415$	$F(2, 2436)=.32$, $p=.867$	$F(4, 2457)=.10$, $p=.983$	$F(2, 1719.28)=.07$, $p=.992$
Emotion x Race x Feature	$F(24, 2457)=1.95$, $p=.004$*	$F(24, 2436)=1.12$, $p=.311$	$F(24, 2457)=.678$, $p=.878$	$F(24, 1719.10)=1.12$, $p=.318$

* for $p<.05$

** for $p<.001$

6.4.2.1. Total Number of First Fixations

For total number of first fixations, there was a main effect of feature, $F(2, 2457)=10.42$, $p<.001$, and an interaction between race, emotion, and feature, $F(24, 2457)=1.95$, $p=.004$, see Figure 6.3. Post-hoc tests using Bonferroni Pairwise Comparisons showed that participants tended to direct their attention first to the nose more than the mouth ($p<.001$) and eyes ($p<.001$); however, the total number of first fixations landing on the mouth and eyes did not differ significantly ($p=1$).

To examine the interaction between race, emotion, and feature, three separate Linear Mixed Models were conducted for African, East Asian, and Western Caucasian faces. For African faces, there was an interaction between emotion and feature, $F(12, 819)=2.17$, $p=.011$. To examine the interaction between emotion and feature, separate Linear Mixed Models were conducted for each of the seven emotions. Participant ID, face ID, and feature were included as factors (dependent variable = total number of first fixations; fixed factor = feature of face). There was a main effect of feature for "disgust," $F(2, 117)=6.91$, $p=.001$, and a trend towards a main effect for "happy," $F(2, 117)=2.63$, $p=.076$, but not for "neutral," $F(2, 117)=2.19$, $p=.116$, "surprise," $F(2, 117)=1.97$, $p=.145$, "fear," $F(2, 117)=.733$, $p=.483$, "anger," $F(2, 117)=.788$, $p=.457$, and "sad," $F(2, 117)=.11$, $p=.895$. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for "disgust" faces, participants first fixated on the nose more than the eyes ($p=.005$) and mouth ($p=.005$); however, the total number of first fixations landing on the eyes and mouth did not differ significantly ($p=1$). For "happy" faces, there was a trend

towards more first fixations on the eyes than the nose (p=.093); however, the total number of first fixations on the mouth did not differ significantly from the eyes (p=.276) and nose (p=1). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and emotion were included as factors (dependent variable = total number of first fixations; fixed factor = emotion). There was a main effect of emotion for the nose, $F(6, 273)=2.38$, p=.029, but not for the eyes, $F(6, 273)=1.49$, p=.183, and mouth, $F(6, 273)=.469$, p=.831. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that there was a trend towards more first fixations landing on the nose for "happy" than "disgust" faces (p=.052).

There were no significant main effects or interactions for East Asian faces. As for Western Caucasian faces, there was a main effect of feature only, $F(2, 819)=9.03$, p<.001. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that participants tended to first fixate on the nose significantly more than mouth (p=.002) and eyes (p<.001) when perceiving Western Caucasian faces; however, the total number of first fixation for the mouth and eyes did not differ significantly (p=1).

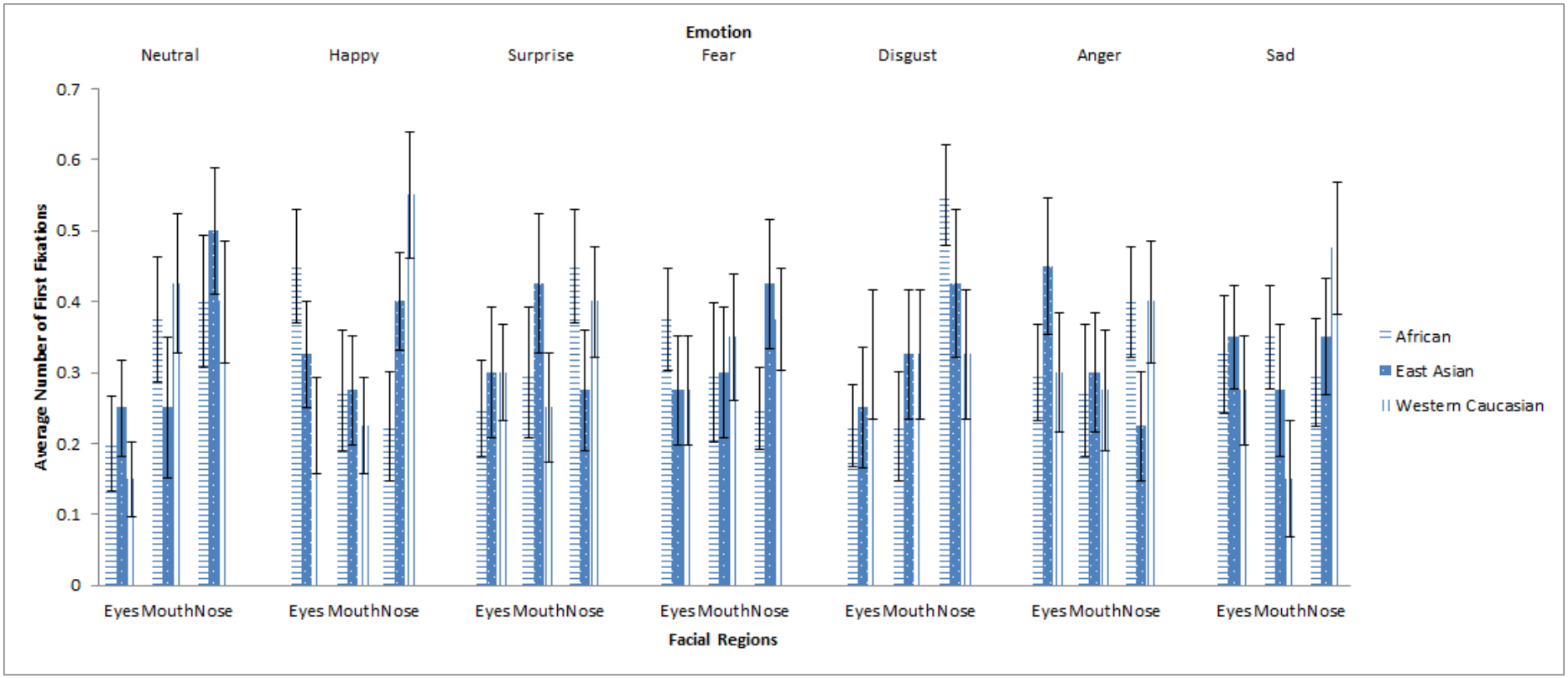


Figure 6.3. Average number of first fixations landing on the eyes, mouth, and nose during the emotion recognition task. Error bars report standard errors of mean. Participants directed their attention first to the nose more than the mouth and eyes.

6.4.2.2. Total Fixation Count

For total fixation count, there was a main effect of feature, $F(2, 2436)=133.34$, $p<.001$, and a trend towards a main effect of race, $F(2, 21)=2.88$, $p=.079$, see Figure 6.4. For feature, post-hoc tests using Bonferroni Pairwise Comparisons revealed participants fixated on the mouth ($p<.001$) and nose ($p<.001$) significantly more frequently than the eyes; however, the total fixation count for the mouth and nose did not differ significantly ($p=.206$). As for race, post-hoc tests using Bonferroni Pairwise Comparisons showed that there was a trend towards participants fixating on East Asian more than African faces ($p=.078$); however the total fixation count for Western Caucasian faces was not significantly different from East Asian ($p=.895$) and African ($p=.596$) faces.

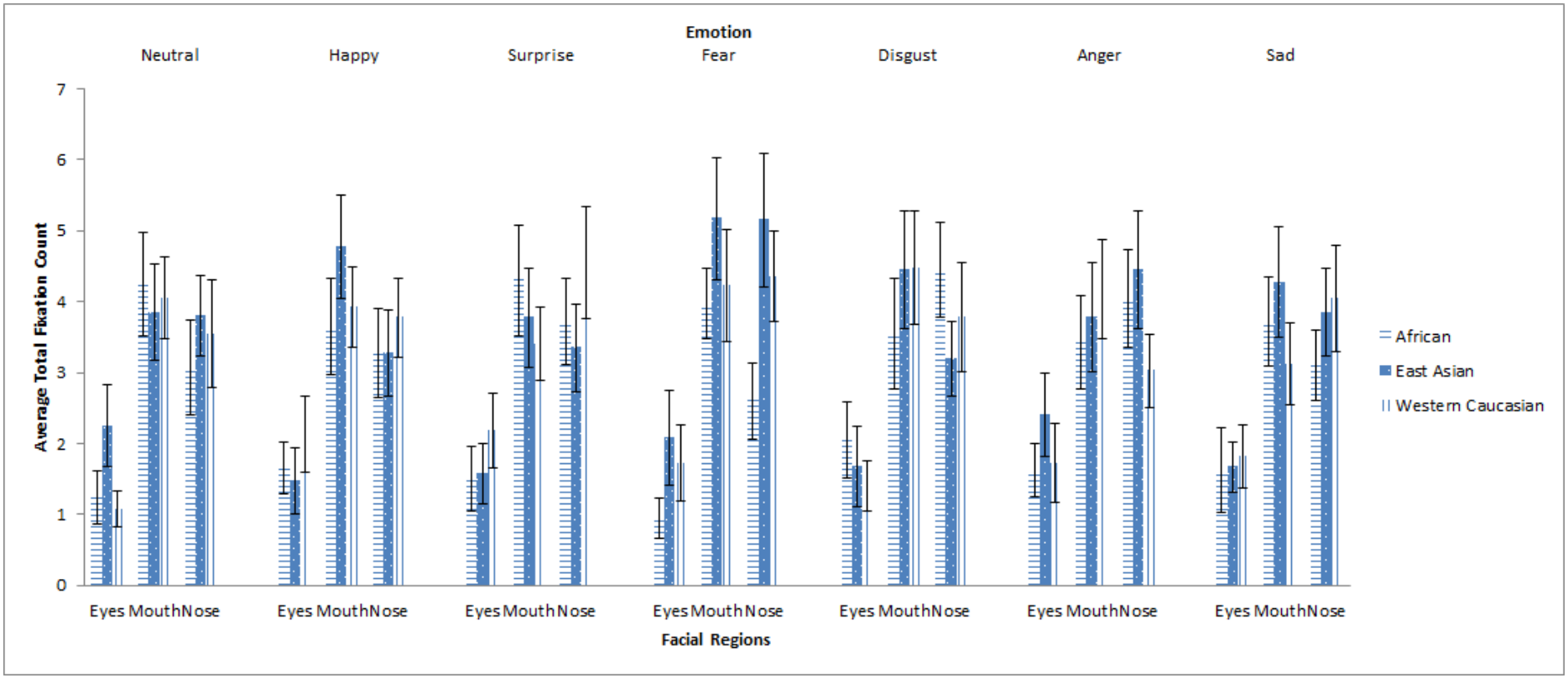


Figure 6.4. Average total fixation count for the eyes, mouth, and nose during the emotion recognition task. Error bars report standard errors of mean.

Participants fixated on the mouth and nose more frequently than the eyes.

6.5. Dynamic Face Recognition and Emotion Recognition Tasks

As presented in Chapter Four, research has revealed that observers' eye movement strategies may be affected by the nature of the task (Yarbus, 1967; Andrews & Coppola, 1999; Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Jack et al., 2009). The comparison of Malaysian Chinese participants' eye tracking data for the static face recognition and emotion recognition tasks has provided further support that participants adapt their looking strategy according to the task at hand (see Chapter Four). To examine whether Malaysian Chinese utilize different eye movement strategies to perceive the same dynamic stimuli when instructed to perform different tasks, eye tracking data for six identical videos that were presented in the current study and the dynamic face recognition study (presented in Chapter Five) were analyzed. The 6 videos of models introducing themselves were evenly distributed among race (i.e. African, East Asian, and Western Caucasian) and gender. In the dynamic face recognition task, the introduction videos were shown in the recognition phase of the experiment and participants were asked to indicate whether they have seen each face before, whereas in the dynamic emotion recognition task, the same videos were used as "neutral" videos and participants were asked to identify which emotions the models were describing.

The procedure for processing and analyzing the eye tracking data was identical to those presented in Chapter Four, see Table 6.5 for a summary of statistics for each dependent variable. For the Linear Mixed Modeling analyses, participant ID nested within task and face ID nested within race were included as a random factor for total fixation count and mean fixation duration. Only participant ID nested within task was included as a random factor for total fixation duration and no random factors

were included for total number of first fixations because the final Hessian matrix was not positive, which casts doubt on the validity of results; however, all results with or without face ID nested within race included as a random factor were identical.

Analyses for total fixation duration revealed a pattern similar to those observed in the total fixation count; therefore, the results for total fixation duration were not discussed in the current chapter (refer to Appendix F for the results of total fixation duration).

Table 6.5. Summary of statistics for each dependent variable in the linear mixed models.

	Total Number of First Fixations	**Total Fixation Count**	**Total Fixation Duration**	**Mean Fixation Duration**
Task	$F(1, 645)=.004$, $p=.949$	$F(1, 35)=1.46$, $p=.235$	$F(1, 35)=.70$, $p=.408$	$F(1, 35.23)=.25$, $p=.622$
Race	$F(1, 3.00)=.38$, $p=.714$	$F(2, 3.02)=.10$, $p=.990$	$F(2, 613)=.12$, $p=.884$	$F(2, 3.01)=.001$, $p=.999$
Feature	$F(2, 645)=.02$, $p=.979$	$F(2, 610)=55.63$, $p<.001$**	$F(2, 613)=59.87$, $p<.001$**	$F(2, 460.20)=10.91$, $p<.001$**
Task x Race	$F(2, 645)=.88$, $p=.415$	$F(2, 610)=1.23$, $p=.294$	$F(2, 613)=.91$, $p=.405$	$F(2, 452.30)=.07$, $p=.930$
Task x Feature	$F(2, 645)=.01$, $p=.993$	$F(2, 610)=8.59$, $p<.001$**	$F(2, 613)=4.93$, $p=.008$*	$F(2, 460)=.23$, $p=.796$
Race x Feature	$F(4, 645)=.73$, $p=.572$	$F(4, 610)=1.61$, $p=.170$	$F(4, 613)=1.32$, $p=.263$	$F(4, 451.87)=.53$, $p=.712$
Task x Race x Feature	$F(4, 645)=.78$, $p=.540$	$F(4, 610)=.25$, $p=.911$	$F(4, 613)=1.69$, $p=.150$	$F(4, 451.59)=2.41$, $p=.049$*

* for $p<.05$

** for $p<.001$

6.5.1. Total Number of First Fixations

For total number of first fixations, there were no main effects or interactions for all factors, see Figure 6.5.

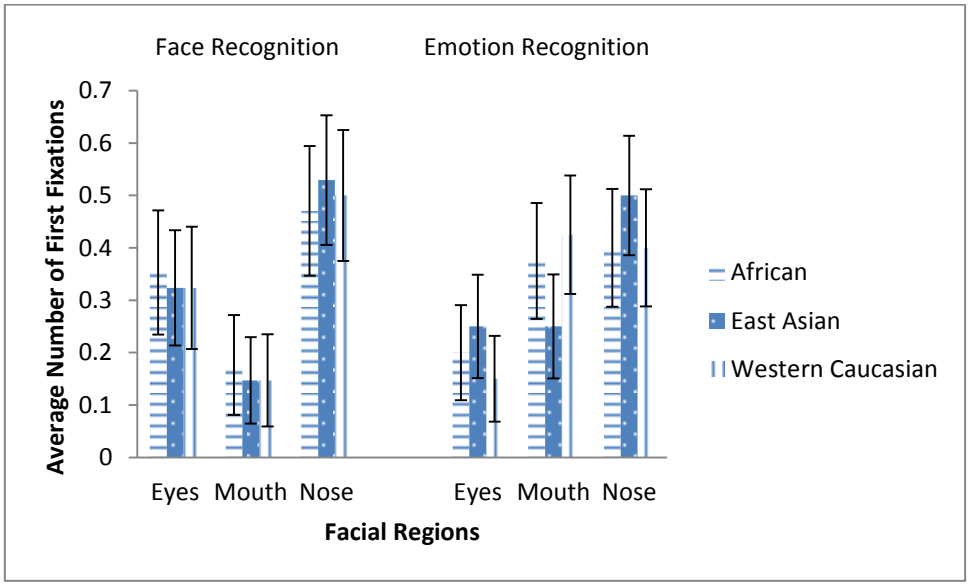


Figure 6.5. Average number of first fixations landing on the eyes, mouth, and nose during the face recognition and emotion recognition tasks. Error bars report standard errors of mean. For first fixations, participants had no preference for the eyes, mouth, or nose during the face recognition and emotion recognition tasks.

6.5.2. Total Fixation Count

For total fixation count, there was a main effect of feature, $F(2, 610)=55.63$, $p<.001$, and an interaction between task and feature, $F(2, 610)=8.59$, $p<.001$, see Figure 6.6. For feature, post-hoc tests using Bonferroni Pairwise Comparisons showed that participants fixated on the nose ($p<.001$) and mouth ($p<.001$) significantly more than the eyes; however the total fixation count for the nose and mouth did not differ significantly ($p=.16$).

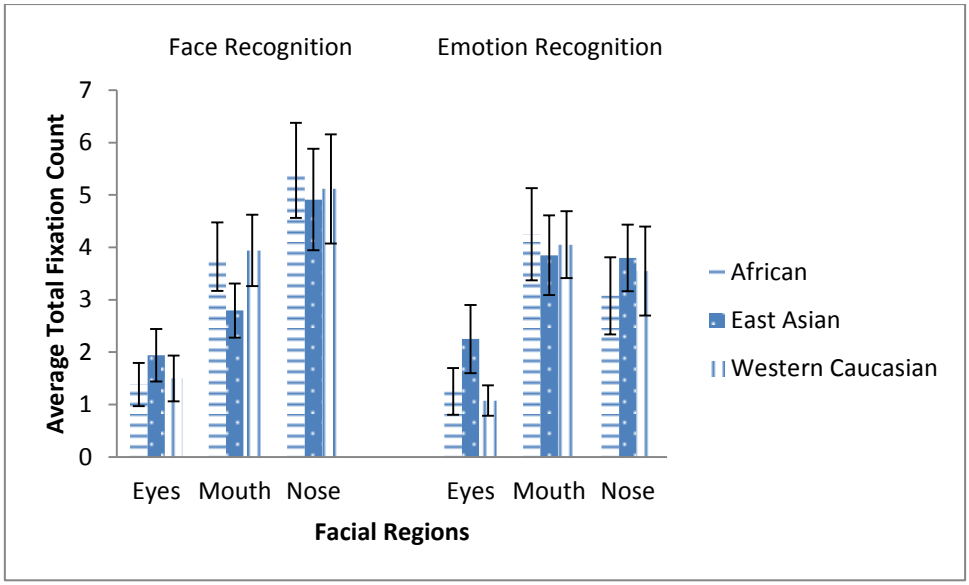


Figure 6.6. Average total fixation count for the eyes, mouth, and nose during the face recognition and emotion recognition tasks. Error bars report standard errors of mean. When watching the same set of introduction videos, participants fixated most frequently on the nose, followed by the mouth then eyes when instructed to recognize identities; however, participants fixated on the mouth and nose more than the eyes when recognizing emotions.

To examine the interaction between task and feature, separate Linear Mixed Models were conducted for each of the two tasks. Participant ID, face ID, and feature were included as factors (dependent variable = total fixation count; fixed factor = feature of face). There was a main effect of feature for both face recognition and emotion recognition tasks, $F(2, 303)=36.75$, $p<.001$ for face recognition task and $F(2, 357)=22.64$, $p<.001$ for emotion recognition task. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for the face recognition task, participants made significantly more fixations on the nose than the mouth ($p<.001$) and eyes ($p<.001$), and the mouth more than the eyes ($p<.001$). For the emotion recognition task, participants fixated on the mouth ($p<.001$) and nose ($p<.001$) significantly more than the eyes; however, the total fixation count for the mouth and nose did not differ significantly ($p=.434$). Further Linear Mixed Models were conducted for each of the

three facial features. Participant ID, face ID, and task were included as factors (dependent variable = total fixation count; fixed factor = task). There was a main effect of task for the nose, $F(1, 220)=11.87$, $p=.001$, but not the eyes, $F(1, 220)=.08$, $p=.777$, and mouth, $F(1, 220)=1.69$, $p=.195$. For the nose, there were significantly more fixations for the face recognition than emotion recognition task ($p=.001$).

6.5.3. Mean Fixation Duration

For mean fixation duration, there was a main effect of feature, $F(2, 460.20)=10.91$, $p<.001$, and an interaction between task, race, and feature, $F(2, 451.59)=2.41$, $p=.049$, see Figure 6.7. For feature, post-hoc tests using Bonferroni Pairwise Comparisons showed that the mean fixation duration for the mouth ($p<.001$) and nose ($p=.001$) was significantly longer than the eyes; however the mean fixation duration for the mouth and nose did not differ significantly ($p=.956$).

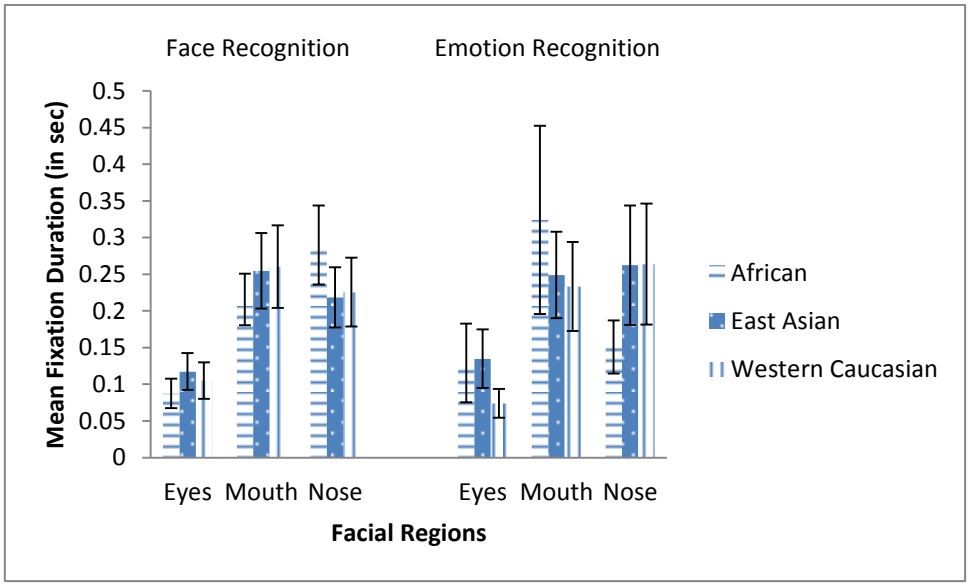


Figure 6.7. Mean fixation duration for the eyes, mouth, and nose during the face recognition and emotion recognition tasks, in seconds. Error bars report standard errors of mean. The mean fixation duration for the mouth and nose was significantly longer than the

eyes for the face recognition task, whereas the mean fixation duration for the eyes, mouth, and nose was not significantly different from each other for the emotion recognition task.

To examine the interaction between task, race, and feature, two separate Linear Mixed Models were conducted for the face recognition and emotion recognition tasks. For the face recognition task, there was a main effect of feature only, $F(2, 234.42)=13.55$, $p<.001$. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for the face recognition task, the mean fixation duration for the mouth ($p<.001$) and nose ($p<.001$) was significantly longer than the eyes; however, the mean fixation duration for the mouth and nose did not differ significantly ($p=1$). No main effects or interactions were found for the emotion recognition task.

6.6. Discussion

The current study investigated Malaysian Chinese participants' eye movement strategy and categorization sensitivity in recognizing African, East Asian, and Western Caucasian facial expressions of emotion. Rather than showing posed expressions, which have been shown to differ from spontaneous expressions in terms of degree of irregularity and speed of onset and offset of expressions (Hess & Kleck, 1990), the current study aimed to develop more naturalistic stimuli that closely represent real life interactions by utilizing muted videos of models describing an emotional experience. This captures a more realistic scenario people may encounter in everyday life, with models possibly de-intensifying or masking their actual emotions while narrating emotional experiences that they may not be comfortable sharing about openly (Ekman, 1970). The primary findings were that 1) Malaysian Chinese participants were best at categorizing "happy" while there was a trend towards "fear"

being recognized worse than “anger” and “disgust”, 2) attention was directed first to the nose more than the mouth and eyes, and 3) participants fixated on the mouth and nose more frequently than the eyes.

Malaysian Chinese participants were able to recognize all emotions described by African, East Asian, and Western Caucasian models above chance level. The behavioral results showed that participants were best at categorizing “happy” for both dynamic (*M*=.80, *SD*=.12) and static conditions (*M*=.96, *SD*=.01; as presented in Chapter Four). However, dissimilar to the results in the static condition which found that participants recognized “fear” worse than all other emotions, the current study showed a trend towards “fear” being recognized worse than “anger” and “disgust” only in the dynamic condition. Additionally, “sad,” which was the second best recognized expression in the static condition (*M*=.92, *SD*=.05), became the second worst recognized emotion in the dynamic condition (*M*=.66, *SD*=.15). Previous studies have found quicker onset and offset, and more pauses and changes in intensity for deliberate than spontaneous dynamic facial expressions, indicating observable differences in posed and natural facial expressions of emotion (Hess & Kleck, 1990). In addition, Ekman (1970) has proposed that different evoking stimuli, display rules, and behavioral consequences may govern one’s expressions of emotion in different social settings. Therefore, it is plausible that models may have either dramatized their posed expression of “sadness” making it easier for participants to recognize the emotion in the static condition, de-intensified their narration of a sad experience making it harder for participants to recognize the emotion in the dynamic condition, or both.

Eye tracking results showed that participants first fixated on the center of the face (i.e. in the nose area) more than the mouth and eyes, possibly to gain a more

holistic view of the face at first glance. When the interaction between emotion and feature was investigated separately by race, results indicated that participants tended to first fixate on the nose for "happy" more frequently than "disgust" emotions described by African models, first fixations landing on the eyes, mouth, and nose did not differ significantly for East Asian emotions, and participants followed the same pattern (with more first fixations landing on the nose than the mouth and eyes) when viewing Western Caucasian videos. This may suggest that although participants generally fixate first on the nose to gain a holistic view of the face, the eyes and mouth may be more expressive when African and East Asian models describe different emotional experiences, resulting in a diversion of initial attention.

Unlike the strategy employed in the static emotion recognition task (in which participants focused more on the eyes, followed by the nose then mouth; reported in Chapter Four), Malaysian Chinese participants fixated on the mouth and nose more frequently than the eyes when viewing dynamic videos of all races of faces. There was also a trend towards more fixations for East Asian than African faces. Participants may have shifted their attention away from the eyes and towards the mouth and nose because the mouth has been identified as the most expressive region of the face, providing more useful cues for identifying one's emotion (Yuki, Maddux, & Masuda, 2007; Blais et al., 2012). A previous eye tracking study has also suggested that the focus on the mouth is not affected by participants' attempt to lip read, but rather the degree of task-relevant information a feature provides (Vo, Smith, Mital, & Henderson, 2012). Focusing on the center of the face (in the nose area) may also be an optimal strategy for participants to observe the models' facial and head movements from a holistic view. Additionally, directing attention towards the mouth is a pattern in line with previous face processing studies utilizing naturalistic videos of faces,

which suggests that although focusing on the eyes may be advantageous for perceiving static stimuli, participants consistently preferred the lower region of the face (i.e. the nose and mouth area) when presented with naturalistic dynamic stimuli (Pillai, Sheppard, & Mitchell, 2012; Chapter 5).

Besides analyzing the current dynamic emotion recognition results, selected data from the current study were compared with data collected from the dynamic face recognition task (presented in Chapter Five). Participants in both studies were presented with 6 identical videos of models introducing themselves, but were required to perform different tasks (i.e. participants were required to identify whether they have seen the faces before in the learning phase of the dynamic face recognition task versus to categorize the emotions that the models are describing in the dynamic emotion recognition task). This gave rise to the opportunity to directly compare eye movements for these particular videos when viewed in two differing tasks. Eye tracking data (i.e. total number of first fixations, total fixation count, total fixation duration, and mean fixation duration) for the 6 videos were analyzed to examine whether participants employ similar or different eye movement strategies when given different set of instructions before viewing the same videos. Separate analyses on the dynamic face recognition and emotion recognition tasks revealed that participants tended to fixate first on the nose when performing both tasks; however, analyses comparing both tasks showed that participants had no preference for looking first at the eyes, mouth, or nose for all three races of faces during both tasks. This indicates that participants tended to direct their attention first to the eyes, mouth, and nose in a similar manner when viewing the six neutral introduction videos.

Analyses on total fixation count revealed that participants employed different fixation patterns for the same stimuli when instructed to perform different tasks. For

the face recognition task, participants fixated significantly more on the nose, followed by the mouth then eyes, whereas participants fixated on the mouth and nose more frequently than the eyes for the emotion recognition task. Analyses for feature also showed that participants fixated significantly more on the nose for the face recognition than emotion recognition task.

Because the eye tracking data examined in the current comparison analyses were selected from two groups of participants who received two different sets of instructions before viewing 6 identical videos, these findings provided more direct evidence that participants do alter their strategy based on participants' goals and aims. Fixations on the mouth increased in the dynamic emotion recognition task, as compared to the dynamic face recognition task, possibly because the mouth contained more diagnostic information that can be used to identify the emotions described by models, whereas focusing on the nose area to gain a more holistic view of models' facial and head movements may be a more optimal strategy to recognize identities (Yuki et al., 2007; Hill & Johnston, 2001).

Lastly, analyses for mean fixation duration revealed that participants tended to fixate on the mouth and nose significantly longer than the eyes for the face recognition task but showed no preference for the eyes, mouth, or nose for the emotion recognition task. This result indicates that although participants looked at the nose more frequently than the mouth during the face recognition task, participants made longer fixations on the mouth than the nose, resulting in similar mean fixation duration for the two features. As for the emotion recognition task, participants made longer but fewer fixations on the eyes and shorter but more fixations on the nose and mouth, resulting in similar mean fixation duration for all three features. This result implies that the eyes may be an equally informative feature that provides crucial

information about one's emotion, which is a pattern in line with previous studies showing that East Asian observers tended to rely on the eye region to interpret facial expressions of emotion as it provides useful cue to identify one's true emotion (Yuki et al., 2007; Ekman, 1992).

Analyses comparing eye tracking data for the dynamic emotion recognition and face recognition tasks revealed that Malaysian Chinese participants altered their fixation pattern when instructed to perform different tasks on an identical set of stimuli, which provides direct support that participants adopt different eye movement strategies according to their goals and aims to enhance their behavioral performance.

Chapter Seven

General Discussion and Conclusions

7.1. Overview of Research

Until recently, it was assumed that the cognitive processes involved in face processing are similar across cultures. More recently, however, eye tracking studies have revealed differences, with East Asian participants employing a more holistic looking pattern, and Western Caucasian participants a more analytic looking pattern when performing various cognitive and perceptual tasks, including face perception.

This thesis investigated Malaysian Chinese participants' eye movement strategy and sensitivity in recognizing static and dynamic identities and emotions. Previous eye tracking studies involving East Asian and Western Caucasian observers have revealed distinct Eastern and Western eye movement strategies when identifying images of faces and facial expressions of emotion (Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Jack, Blais, Scheepers, Schyns, & Caldara, 2009). More specifically, studies in East Asian populations have shown that while participants focused on the center of the face in face recognition tasks, participants adapted their strategy when asked to identify emotions rather than identity—they focused more on the eyes, suggesting that the eyes may contain more salient information pertaining to emotional state in Eastern cultures. Studies in Western Caucasian populations, on the other hand, have found that participants look evenly across the face by moving between the eyes and mouth to identify both identity and emotions. Additionally, face recognition studies investigating the effects of cultural

environment on populations, such as a study that recruited British born Chinese, revealed that exposure to and familiarity with another culture may foster a hybrid fixation pattern that enabled observers to recognize both own- and familiar other-race faces equally well (Kelly, Jack, Miellet, De Luca, Foreman, & Caldara, 2011a). Therefore, exploring how participants from unique cultures, such as Malaysia, would perform in face perception tasks will give us insight as to whether observers originating from Eastern countries that are greatly influenced by Western culture would employ eye movement strategies more alike with East Asians or more in line with those of British born Chinese. To produce results comparable to previous eye tracking studies involving different populations (e.g. East Asians, Western Caucasians, and British born Chinese), Chapter Three and Chapter Four utilized static images of faces and facial expressions of emotion to examine Malaysian Chinese participants' behavioral and eye tracking performance.

Even though participants showed high performance level in laboratory settings (e.g. face recognition studies), accurate recognition abilities do not always seem to be replicable in real life situations, such as eyewitness testimony. As discussed in Chapter One, the incongruence between recognition performance in laboratory studies and real life settings may be a result of the artificial nature of laboratory studies. Although utilizing static images provides greater control of the development and presentation of stimuli, performance in studies using images may not reflect real life situations particularly well. For instance, images only provide a two-dimensional view of the face; however, studies have shown that rotating faces in depth may impair participants' recognition accuracy, suggesting that a change of perspective and presentation method, which would necessarily occur from laboratory studies to real life situations, may reduce

recognition accuracy (Valentin, 1999; Hill, Schyns, & Akamatsu, 1997; O'Toole, Jiang, Roark, & Abdi, 2006). Moreover, real faces are dynamic, which would provide additional information about one's external and internal state through facial motions or head movements. The lack of dynamic information in static images may produce results different from those observed in studies utilizing naturalistic stimuli. Because of these methodological considerations, naturalistic videos of models introducing themselves and describing different emotional experiences were collected for Chapter Five and Chapter Six respectively. These two studies examined whether Malaysian Chinese participants adopt different eye movement strategies when viewing stimuli which more accurately represent a natural social situation and participants' sensitivity in recognizing identities and emotions from moving stimuli. Previous eye tracking studies have found that East Asian younger children and adults tended to fixate longer on the eyes whereas Western Caucasian younger children and adults directed their attention more towards the mouth when viewing dynamic stimuli of models presented in different angles with different eye gaze and movements in the mouth (Senju, Vernetti, Kikuchi, Akechi, & Hasegawa, 2013a; Senju, Vernetti, Kikuchi, Akechi, Hasegawa, & Johnson, 2013b). However, eye tracking results from Chapter Five and Chapter Six may not be directly comparable with previous eye tracking studies because of the different nature of the task (i.e. face recognition or emotion recognition versus passive viewing) and stimuli (i.e. models talking naturally about themselves or an emotional experience versus posing between neutral and happy expressions).

7.2. Summary and Discussion of Results

7.2.1. Chapter Three: Static Face Recognition

Chapter Three explored the effects of exposure and familiarity with Western culture on Malaysian Chinese participants' recognition abilities and eye movement strategy by requiring participants to perform a face recognition task using static images of African, East Asian, and Western Caucasian faces. Behavioral results showed that Malaysian Chinese participants performed equally well at recognizing East Asian and Western Caucasian faces, but less well at recognizing African faces.

When perceiving all three races of faces, participants tended to fixate first on the nose more often than the eyes, followed by the mouth. For the learning task, a majority of participants instinctively use a holistic approach to make first impressions or subjective judgments (i.e. the attractiveness) about a face by fixating first on the nose. However, participants are more likely to rely on the individuating feature of the face when recalling identities in the recognition task, with more first fixations directed towards the eyes.

Participants employed a fusion of Eastern and Western eye movement strategies when perceiving all races of faces, with most fixations clustered around the eyes, followed by the nose then mouth. There were more fixations on the mouth for African and East Asian than Western Caucasian faces. The diversity of the Malaysian culture may play a large role in shaping the distinctive fixation pattern that Malaysian Chinese participants adopt. The influence from the West, as reflected in Malaysia's history (Advameg, 2011; Gaudart, 1987) and media (Davies, 2011; Epstein, 2011), may have shaped Malaysian Chinese observers' fixation pattern to optimally recognize the diverse

group of people they frequently encounter. Participants fixated more on the mouth for African faces, a finding in line with previous findings that different facial features are relevant to different races of faces; more specifically, the lower facial features have been found to be the diagnostic features for African faces (Ellis, Deregowski, & Shepherd, 1975; Hills and Lewis, 2006).

Lastly, the mean fixation duration showed that Malaysian Chinese participants made fixations of longer duration on the eyes and mouth than the nose, indicating that participants had fewer but longer fixations on the mouth and more short fixations on the nose. In line with the findings for total fixation duration, participants made fixations of longer duration for African than East Asian and Western Caucasian faces, especially on the mouth and nose. The lower facial features of African faces not only rendered a greater fixation count but also longer total and mean fixation duration, providing stronger evidence that the lower facial features may be especially informative for identifying and making subjective judgments of African faces. For East Asian faces, participants made fixations of significantly longer duration on the mouth than the eyes, followed by the nose, whereas the mean fixation duration on the eyes was significantly longer than the nose for Western Caucasian faces, suggesting that the eyes may convey more individuating information (e.g. eye color) for Western Caucasian than East Asian and African faces, which tend to have rather uniform eye color (Ellis et al., 1975).

7.2.2. Chapter Four: Static Emotion Recognition

Chapter Four aimed to examine how Malaysian Chinese participants would perform when categorizing static images of East Asian and Western Caucasian facial

expressions of emotion. Behavioral results showed that Malaysian Chinese participants were best at categorizing "happy" whereas "fear" was recognized less well than other emotions. Participants were also better at recognizing East Asian than Western Caucasian facial expressions, although this was only true for "happy" and "surprise." It is unlikely that this is due to unfamiliarity with Western Caucasian faces because it was shown that Malaysian Chinese recognized Western Caucasian faces with equal proficiency as East Asian ones in the static face recognition study in Chapter Three. Participants were more proficient in recognizing certain own-race emotions possibly because certain portrayal of emotions is more relevant or familiar to their culture while there may be a discrepancy between the way expressions were posed and perceivers' conception of those emotions for other-race faces (Ekman & Friesen, 1971; Biehl, Matsumoto, Ekman, Hearn, Heider, Kudoh, & Ton, 1997).

When categorizing facial expressions of emotion, Malaysian Chinese observers generally directed their attention first to the nose more often than the eyes and mouth, possibly to gain a more holistic view of the faces.

Malaysian Chinese observers produced most fixations on the eyes, followed by the nose then mouth when perceiving most facial expressions of emotion; however, for "happy," "disgust," and "sad" faces, there were more fixations on the eyes and nose than the mouth. Similar to the results of a previous eye tracking study, participants directed more attention to certain regions of the face for different facial expressions, which suggests that different emotions may be portrayed differently. Therefore, fixating on the features of the face that may be more expressive for certain emotions would provide more diagnostic information to identify the emotions accurately (Guo, 2012).

Lastly, the mean fixation duration showed that participants made fixations of a longer duration on the mouth than the eyes and nose for most facial expressions of emotion, implying that participants made longer but fewer fixations on the mouth and shorter but more fixations on the eyes and nose. Although analyses for the total fixation count indicated that there were more fixations on the eyes and fewer fixations on the mouth, the mean fixation duration was longer for the mouth than the eyes. This suggests that participants may be directing their gaze back to the eyes frequently after looking at other features (example of gaze pattern: eyes, mouth, eyes, nose, eyes, mouth, eyes, nose, eyes, etc.) or that participants' gaze may be moving around a lot within the region (especially because there are two separate areas of interest for the eyes, participants may be looking at each eye each time they look at the eye region; therefore, focusing on the eye region may render more fixations if participants moved back and forth between the left eye and right eye).

7.2.3. Static Face Recognition and Emotion Recognition Tasks

Previous face processing studies have revealed that observers employ different eye movement strategies according to the nature of the task (Yarbus, 1967; Andrews & Coppola, 1999; Blais et al., 2008; Jack et al., 2009). Therefore, further analyses were conducted to compare the eye tracking data for East Asian and Western Caucasian faces when performing the static face recognition and emotion recognition tasks.

Malaysian Chinese participants tended to fixate first on the nose and eyes more than the mouth when performing the static face recognition task, suggesting that the upper region of the face may contain more diagnostic information useful for recognizing

identities. However, there were more first fixations landing on the nose, followed by the eyes then mouth for the emotion recognition task. When perceiving different facial expressions of emotion, participants are unable to anticipate whether the most expressive region of the face would be located in the eyes or mouth area; hence, fixating first on the center of the face may be an optimal strategy to gain a more holistic view of a facial expression. Additionally, results showed that participants directed their attention first to the eyes for East Asian faces significantly more than Western Caucasian faces, a result in line with previous findings that East Asian observers learn to rely on the eye region for more accurate information about an East Asian's true emotion (Yuki, Maddux, & Masuda, 2007).

Malaysian Chinese participants appeared to have employed a similar eye movement strategy (with more fixations on the eyes than the nose, followed by the mouth) when perceiving both East Asian and Western Caucasian faces during both face recognition and emotion recognition tasks; however, analyses comparing the total fixation count for both tasks revealed that participants tended to look less at the eyes and more at the nose and mouth in the emotion recognition task, compared to the face recognition task. This change of strategy is in line with previous research which found a change of eye movement patterns in East Asian observers for these two types of face processing tasks; more specifically, East Asians employed a holistic strategy by fixating mostly on the nose to recognize identities but focused on the eyes to recognize identities because they bear diagnostic information of an individual's true emotion (Blais et al., 2008; Jack et al., 2009; Yuki et al., 2007). Malaysian Chinese participants adopted a mixture of Eastern and Western strategies in the static face recognition task to better

recognize own- and other-race faces. Even though participants shifted their attention towards the nose and mouth in the emotion recognition task, which suggests a move towards a Western strategy, fixations were directed mainly towards the eyes, a strategy that may have enabled Malaysian Chinese to better recognize certain own- than other-race facial expressions. Previous studies have suggested that observers learn to rely on different facial cues because of the cultural differences in how the emotions are expressed (Yuki et al., 2007). Thus, in a culture where the control of emotion to maintain harmonious relationships with others is valued, East Asian observers may have learned that fixating more on the eyes, where the muscles are more difficult to control than the muscles around the mouth, may reveal more diagnostic cue to identify East Asians' true emotions (Friesen, 1972; Ekman, 1992).

Finally, the mean fixation duration on the mouth and eyes were significantly longer than the nose for the static face recognition task, while the mean fixation duration on the mouth was significantly longer than the eyes, followed by the nose for the static emotion recognition task. This indicated that participants tended to direct more attention to the mouth and eyes when asked to recognize identities whereas images of facial expressions rendered longer but fewer fixations on the mouth and shorter but more fixation on the eyes than nose possibly because the mouth has been identified as the most expressive region of the face, providing useful cues for observers to categorize facial expressions accurately (Yuki et al., 2007).

7.2.4. Chapter Five: Dynamic Face Recognition

Chapter Five aimed to use naturalistic stimuli that more accurately represent a typical social situation (i.e. muted videos of models introducing themselves) to investigate how Malaysian Chinese participants would perform in a dynamic face recognition task. Behavioral results revealed an equal ability in recognizing African, East Asian and Western Caucasian faces, suggesting that dynamic stimuli may, perhaps, reduce the other-race effect by providing more diagnostic information about a face that is useful for recognizing identities.

The total number of first fixations on the nose was significantly more than the eyes, followed by the mouth for the learning and recognition tasks, a pattern similar to those observed in the static face recognition task. This suggests that participants consistently fixate first on the center of the face to gain a more holistic view at first glance despite a change in stimuli.

For both learning and recognition tasks, the total fixation count for the nose was significantly more than the mouth, followed by the eyes for all three races of faces. As compared with the previous static face recognition study that utilized an identical procedure, Malaysian Chinese participants employed different strategies to perceive static images and dynamic videos; participants focused on the eyes, followed by the nose then mouth when viewing images, whereas attention was directed towards the nose, followed by the mouth then eyes when viewing videos of faces. Focusing on the eyes may be an optimal strategy to perceive images as previous studies have shown that the eyes provide differentiating cues for recognizing faces in static stimuli (Ellis et al., 1975); however, when perceiving videos, participants may focus on the center of the face to gain a more

holistic view of facial and head movements that have been shown to provide useful cues in recognizing identities (Hill & Johnston, 2001). In addition, the increase in fixations for the mouth in the dynamic task is in line with previous eye tracking study investigating participants' judgments about scenarios based on recorded videos of models' natural reactions (Pillai, Sheppard, & Mitchell, 2012). Although it is possible that participants may be attracted to the mouth based on visual saliency, a previous eye tracking study has suggested that fixations are modulated by the degree of task-relevant information a feature provides, indicating that the mouth may provide useful cues for identifying faces (Vo, Smith, Mital, & Henderson, 2012). It is unlikely that participants were trying to lip-read the muted videos as study has found that participants tended to fixate less on the mouth when viewing muted videos, as compared to videos with audio, when judging the likeability of naturalistic videos (Vo et al., 2012). These findings suggest that although the eyes may provide useful cues in recognizing static images of faces, focusing on the lower regions of the face (i.e. the nose and mouth area) would allow observers to gain a more holistic view of the models' facial and head movements and may provide important cues for recognizing dynamic faces.

Lastly, the mean fixation duration revealed that participants made longer fixations on the mouth than the nose, followed by the eyes, implying that participants made fewer but longer fixations on the mouth. This finding suggests that although participants tended to fixate more frequently on the center of the face to gain a more holistic view of models' facial and head movements, participants fixated longer on the mouth to process cues advantageous for recognizing dynamic faces (Pillai et al., 2012; Vo et al., 2012).

7.2.5. Chapter Six: Dynamic Emotion Recognition

The final study examined Malaysian Chinese participants' sensitivity and eye movement strategy when perceiving muted videos of models describing emotional experiences. The behavioral results revealed that participants were able to recognize all emotions described by African, East Asian, and Western Caucasian models above chance level. Participants were best at categorizing "happy" for both static and dynamic emotion recognition tasks. However, unlike the static emotion recognition task which found that participants recognized "fear" worse than all other emotions, the current study showed a trend towards "fear" being recognized worse than "anger" and "disgust" only. The different behavioral performance in the static and dynamic emotion recognition tasks may be due to the differences in the portrayal of posed and natural expressions and the different evoking stimuli, display rules, and behavioral consequences that may govern one's expressions of emotion in different social settings, providing evidence that future studies should focus more on the development of naturalistic stimuli that more accurately represent a typical social situation to produce more ecologically valid results (Hess & Kleck, 1990; Ekman, 1970).

Participants fixated first on the nose more often than the mouth and eyes, possibly to gain a more holistic view of the face at first glance. Although participants generally fixated first on the center of the face, participant tended to divert their initial attention to the eyes and mouth for African and East Asian faces, possibly because different features may be more expressive for certain emotions when African and East Asian models describe different emotional experiences.

Malaysian Chinese participants fixated on the mouth and nose more frequently than the eyes when viewing dynamic videos of all races of faces, a strategy different from those employed in the static emotion recognition task (in which participants focused more on the eyes, followed by the nose then mouth). Although focusing on the eyes may be advantageous for perceiving static stimuli, participants consistently preferred the lower region of the face (i.e. the nose and mouth area) when perceiving dynamic stimuli, a pattern in line with previous studies utilizing naturalistic videos of faces (Pillai et al., 2012; Chapter 5). A previous eye tracking study has found that participants tended to direct their attention towards facial features that provide cues useful for performing the task at hand (Vo et al., 2012). Therefore, participants may have shifted their attention away from the eyes and towards the mouth and nose possibly because the mouth, which has been identified as the most expressive region of the face, would provide more important cues for identifying emotions (Yuki et al., 2007; Blais et al., 2012). Additionally, focusing on the center of the face (in the nose area) may provide a more holistic view for participants to observe the models' facial and head movements.

7.2.6. Dynamic Face Recognition and Emotion Recognition Tasks

Comparison of eye tracking data from the static face recognition and emotion recognition tasks have revealed that participants adapt their eye movement strategy according to the task at hand. To investigate whether Malaysian Chinese employ different eye movement strategies to perceive the same dynamic stimuli when instructed to perform different tasks (i.e. participants were required to identify whether they have seen the faces before in the dynamic face recognition task versus to categorize the emotions that the models are describing in the dynamic emotion recognition task), eye tracking

data for six identical videos that were presented in the dynamic emotion recognition study and the dynamic face recognition study were analyzed. Separate analyses on the dynamic face recognition and emotion recognition tasks revealed that participants tended to fixate first on the nose when performing both tasks; however, analyses comparing both tasks showed that participants had no preference for looking first at the eyes, mouth, or nose for all three races of faces during both tasks. This suggests that participants tended to direct their attention first to the eyes, mouth, and nose in a similar manner when viewing the six neutral introduction videos.

Analyses on the total fixation count revealed that participants utilized different strategies for the same stimuli when instructed to perform the two different tasks. Participants fixated significantly more on the nose, followed by the mouth then eyes when performing the dynamic face recognition task, whereas for the emotion recognition task, participants fixated on the mouth and nose more frequently than the eyes. There were more fixations on the mouth in the dynamic emotion recognition task, as compared to the dynamic face recognition task, possibly because the mouth contained more diagnostic information that can be used to identify the emotions described by models; on the other hand, focusing on the nose area to gain a more holistic view of models' facial and head movements may be a more optimal strategy for recognizing identities (Yuki et al., 2007; Hill & Johnston, 2001).

Finally, the mean fixation duration revealed that participants fixated on the mouth and nose longer than the eyes for the face recognition task but showed no preference for the eyes, mouth, or nose for the emotion recognition task. Although participants looked at the nose more frequently than the mouth during the face recognition task, participants

made longer fixations on the mouth than the nose, resulting in similar mean fixation duration for the two features. As for the emotion recognition task, participants made longer but fewer fixations on the eyes and shorter but more fixations on the nose and mouth, resulting in similar mean fixation duration for all three features. This result indicates that the eyes may be an equally informative feature that provides crucial information about one's emotion, which is consistent with previous findings that East Asian observers rely on the eye region to interpret facial expressions of emotion as it provides useful cue to identify one's true emotion (Yuki et al., 2007; Ekman, 1992).

7.3. Implications of Results

Although many aspects of human cognition and perception have long been assumed to be identical across cultures, cross-cultural studies provided evidence that human thought and visual processes are heavily influenced by the social environment and social practice of individual cultures. As presented in Chapter Two, the Eastern collectivist culture, with its focus on the relationships between individuals within groups, may engender cognitions that are more sensitive to the relationship between objects and the wider context of scenes; on the other hand, the Western individualist culture, with its emphasis on the value of individuals as independent entities, may require a set of cognitive mechanisms that cause one to be more sensitive to the individual objects within a particular scene (Nisbett & Miyamoto, 2005). Moreover, the distinct attentional patterns that East Asian and Western Caucasian participants employ may be fostered by the social practices of their respective cultures. For instance, studies observing mother-and-infant interaction during play found that Japanese mothers attended to the relationship between

an object and its context by using objects to engage their child in social routines while American mothers focused more on the focal object by labelling objects frequently (Fernald & Morikawa, 1993). Attentional differences between Eastern and Western cultures can also be seen in maternal speech and toddlers' use of language (Tardif, Gelman, & Xu, 1999). More specifically, Mandarin-speaking mothers produced more verbs whereas English-speaking mothers used more nouns when communicating with their toddlers. Assessment of the proportion of nouns and verbs also showed the same language differences in toddlers.

The contrasting holistic-analytic styles associated with Eastern and Western cultures are reflected in varying thought and visual processes in both cultures, such as correspondence bias, visual illusory tasks, scene perception, and may possibly extend to the eye movement strategies that East Asian and Western Caucasian participants employ in face perception tasks like face recognition and emotion recognition (Blais et al., 2008; Jack et al., 2009). Research exploring correspondence bias has found that East Asian participants are more likely than Western Caucasian participants to explain and predict behaviors in terms of situational factors; East Asians were more likely to believe that personality may differ according to situational circumstances while Western Caucasians believed personality traits do not vary by situation (Choi & Nisbett, 1998; Norenzayan, Choi, & Nisbett, 2002). Research revealed that East Asian and Western Caucasian participants not only differed in the way they conceptualize information, but also the way they perceive visual fields; more specifically, East Asians tended to integrate the focal object with its context while Western Caucasians focused more on the object alone. As compared with Western Caucasians, East Asian participants' dependence on contextual

information resulted in poorer performance in visual illusory tasks that required strong focus on the focal object while ignoring the context, such as the rod and frame illusion and visual size perception task (Ji, Peng, & Nisbett, 2000; Doherty, Tsuji, & Phillips, 2008). Scene perception studies found similar patterns, with East Asian participants performing poorer than Western Caucasians when recognizing focal object paired with novel backgrounds (Masuda & Nisbett, 2001). A further eye tracking study revealed distinct fixation patterns in line with previous behavioral results; more specifically, East Asians fixated on the focal object and background for similar duration while Western Caucasians fixated quicker and longer at the focal object, resulting in better recognition of the focal object even when paired with novel backgrounds (Chua, Boland, & Nisbett, 2005).

Recent face processing studies utilizing eye tracking methodologies revealed systematic cross-cultural differences in the way East Asian and Western Caucasian participants perceive faces in face recognition and emotion recognition tasks (Blais et al., 2008; Jack et al., 2009). When recognizing identities, East Asian participants focused more on the center of the face, in the nose area, which may indicate a holistic style of processing; however, participants adapted their strategy by focusing more on the eyes when identifying emotions, suggesting that the eyes may contain diagnostic information pertaining to emotional state in Eastern cultures. On the other hand, Western Caucasian participants looked evenly across the face by moving between the eyes and mouth when performing both face recognition and emotion recognition tasks, suggesting a more analytical processing of individual features.

The distinct Eastern and Western eye movement strategies employed in both face processing tasks may be an extension of the contrasting holistic-analytic processing styles associated with the respective cultures as identified in various cognitive and perceptual tasks. However, studies involving Asian participants who are more familiar with the Western culture (e.g. British born Chinese; Kelly et al., 2011a) showed that participants tended to adopt a mixture of both Eastern and Western eye movement strategies when recognizing faces, suggesting that there may be more to cross-cultural differences than a simple Eastern-Western distinction. Culture alone may not fully explain the diversity in participants' eye movement strategies; exposure to and familiarity with the diagnostic features of other-race faces encountered during development may also affect one's looking pattern.

Research has provided evidence that features play a major role in face recognition and different features of the face affect recognition performance in different ways. For instance, internal features were found to be more useful for the recognition of familiar faces whereas both internal and external features were equally useful for the recognition of unfamiliar faces (Ellis, Shepherd, & Davies, 1979). A previous study has also found that participants were most aware of changes in the eyes, followed by the mouth then nose when identifying a target face from a series of distractor faces, indicating that the eyes are the most important features for face recognition (Davies, Ellis, & Shepherd, 1977).

Research suggests that the features of the faces themselves may be partly responsible for driving the cross-cultural differences in Eastern and Western looking strategies as people tended to use different facial features to describe and recognize faces

of different races (Ellis et al., 1975; Hills & Lewis, 2006). A cross-cultural study examining participants' verbal descriptions of faces found that participants from different racial groups used different facial features to describe faces (Ellis et al., 1975). More specifically, Caucasians were more likely to describe features like eye color, while Africans focused on the lower facial features such as lips and noses. In conjunction with this finding, a previous study has found that training participants to focus on the diagnostic features associated with a specific race of face subsequently increased participants' recognition accuracy, thereby reducing the other-race effect (Hills & Lewis, 2006). This finding indicates that the other-race effect may be attributable to the use of looking strategies, learned during development, that are optimized for recognizing familiar own-race faces, but which are suboptimal when applied to unfamiliar other-race faces. This provides evidence that the face processing style and associated looking strategy that observers employ may be determined by the diagnostic features of the faces encountered during development, and that the looking strategy adopted may reflect the variance in facial appearance in the local population.

The other-race effect is a robust phenomenon that has been found to emerge as early as six months old (Kelly, Quinn, Slater, Lee, Ge, & Pascalis, 2007; Kelly, Liu, Lee, Quinn, Pascalis, Slater, & Ge, 2009); however, research has shown that the other-race effect could be reduced by various methods, such as feature-based training, individuation and categorization training, or familiarization through exposure (Hills & Lewis, 2006; Tanaka & Pierce, 2009; Hancock & Rhodes, 2008). A previous study found that training adults to differentiate other-race faces at an individual level and categorical level improved participants' recognition abilities; moreover, participants who received the

individuation training improved more than participants who received the categorization training (Tanaka & Pierce, 2009). These findings suggest that adequate training at an individual level allowed observers to recognize other-race faces better by encoding features of the face, whereas training at a basic level of race shifted observers' attention to race- rather than individual-specific features. Additionally, a study investigating the relationship between contact level, configural coding, and recognition performance has found that higher levels of self-reported contact with other-race faces was associated with a reduced other-race effect in recognition performance and configural coding for both Chinese and Caucasian participants (Hancock & Rhodes, 2008). Even though both quantity and quality of contact were found to be negatively correlated with the other-race effect, the authors proposed that high quantity of contact may not necessarily reduce the other-race effect as it does not require observers to differentiate other-race faces. Instead, quality interaction would allow observers to learn the diagnostic features that individuate faces of other-races, which subsequently reduces the other-race effect (Hills & Lewis, 2006).

Moreover, research has found that the cognitive mechanisms involved in face recognition show plasticity during childhood; more specifically, familiarity with and exposure to other-race faces can reverse the other-race effect (Sangrigoli, Pallier, Argenti, Ventureyra, & De Schonen, 2005; Heron-Delaney, Anzures, Herbert, Quinn, Slater, Tanaka, Lee, & Pascalis, 2011). A cross-cultural study found that Korean adults who were adopted by Caucasian families between ages 3 and 9 were better at recognizing Caucasian than Asian faces, a performance similar to those of Caucasian adults, while native Koreans brought up in Korea recognized Asian faces better than Caucasian ones

(Sangrigoli et al., 2005). This finding indicates that adopted Koreans may have lost the advantage to recognize own-race faces due to familiarity with Caucasian faces, allowing a reversal of the other-race effect even in late childhood. Further study examining the role of perceptual training on the other-race effect found that 9-month-old Caucasian infants who received training on Chinese faces over the course of 3 months were able to recognize other-race faces while infants who did not receive training on Chinese faces could not discriminate between other-race faces (Heron-Delaney et al., 2011). Although the cognitive mechanisms involved in face recognition becomes finely tuned to own-race faces early in life (Kelly, Quinn, Slater, Lee, Ge, & Pascalis, 2007), exposure to other-race faces allow observers to develop a face prototype based on the type of facial input or training received, subsequently reducing the other-race effect.

As compared with East Asian and Western Caucasian participants' behavioral and eye tracking performance in a static face recognition task, results from Chapter Three revealed that perhaps due to the familiarity with and exposure to the Western culture, Malaysian Chinese adopted a combination of both Eastern and Western eye movement strategies to better recognize own- and familiar other-race faces (Blais et al., 2008; Tan et al., 2012). These findings suggest that the cognitive mechanisms involved in face processing show plasticity; learning to use facial cues that have the highest diagnostic value in the local population may improve recognition performance and reduce the other-race effect. The diversity of Malaysian culture, influence from the West (as reflected in Malaysia's history), or familiarity with Western faces through media may have shaped the unique fixation pattern that Malaysian Chinese participants employed to optimally

recognize the diverse group of people they frequently encounter (Advameg, 2011; Gaudart, 1987; Davies, 2011; Epstein, 2011).

Additionally, results from Chapter Four showed that participants were more proficient in recognizing certain own-race emotions possibly because there is a discrepancy between the way expressions were posed and perceivers' conception of those emotions for other-race faces; participants may find certain portrayal of emotions more relevant to their culture and be more proficient in recognizing these expressions (Ekman & Friesen, 1971; Biehl, Matsumoto, Ekman, Hearn, Heider, Kudoh, & Ton, 1997). The cultural differences in the way people express or interpret emotions is a finding more in line with Ekman's (1970) theoretical framework which attempted to reconcile the two extreme ends of the universality versus culture specificity hypotheses; although the facial expressions of the primary emotions are universal to mankind, elements of facial expressions may vary across cultures as a result of factors such as evoking stimuli, display rules, and behavioral consequences of emotion. Further analyses comparing eye tracking data from the static face recognition and emotion recognition tasks showed that Malaysian Chinese participants adapted their looking strategies according to the task at hand, providing further support that eye movements may be altered according to observers' aims and goals.

Although the static face recognition and emotion recognition tasks have provided insight as to how Malaysian Chinese participants, who are more familiar with Western culture, performed in comparison to populations such as East Asians or British born Chinese, research has found that utilizing static stimuli in laboratory studies may have limited ecological validity. Moreover, eye tracking studies have revealed that participants

employed different looking strategies when analyzing dynamic faces, possibly due to a change in the type of stimuli (e.g. static or dynamic stimuli) and/or the nature of the task (e.g. face recognition of static images versus passive viewing of dynamic stimuli). More specifically, East Asian participants showed a preference for the nose when recognizing static images of faces while East Asian younger children and adults tended to direct their attention towards the eyes when perceiving dynamic stimuli of models posing between neutral and happy expressions (Blais et al., 2008; Senju et al., 2013a; Senju et al., 2013b). To identify whether Malaysian Chinese participants altered their eye movement strategies according to the type of stimuli, Chapter Five and Chapter Six presented naturalistic dynamic stimuli that more accurately represent a typical social interaction while employing procedures identical to those utilized in Chapter Three and Chapter Four. Results revealed that Malaysian Chinese adapted different eye movement strategies when performing dynamic face recognition and emotion recognition tasks with identical procedures, suggesting that a change in stimuli does affect participants' strategies, possibly because the diagnostic facial cues differ according to the type of stimuli. To further investigate if participants employed different looking strategies according to the nature of the task, eye tracking data from the dynamic face recognition and emotion recognition tasks were analyzed. Results continued to provide support that participants adapt different eye movement strategies when instructed to perform different tasks (even when presented with identical stimuli).

Previous dynamic face recognition studies have found better recognition performance for familiar faces under non-optimal viewing conditions (Knight & Johnston, 1997; Lander & Bruce, 2000; Lander, Christie, & Bruce, 1999). In Chapter Five, results

showed that Malaysian Chinese participants were able to recognize dynamic own- and other-race faces equally well, at above chance level, when unfamiliar faces were presented under optimal conditions. As suggested by the representation enhancement hypothesis, facial motions may have enhanced participants' encoding of the invariant structure of each face by providing additional structures of the face. Participants may have also learned the models' dynamic facial signatures in the learning phase of the study, which would allow participants to better recognize identities based on facial motions which reflect characteristics that are specific to an individual, supporting the supplemental information hypothesis (O'Toole et al., 2002; O'Toole et al., 2006; Knappmeyer et al., 2003). However, participants did not perform significantly better in the dynamic face recognition task than the static face recognition task possibly because the dynamic task was more difficult. Participants learned and recognized identical images when performing the static face recognition task while in the dynamic task, different segments of each model's video were presented in the learning and recognition phases. Moreover, previous studies have suggested that facial movements may prove challenging to observers as movements require constant examination and interpretation of ample visual cues that may alter the facial structure that is important for recognizing unfamiliar faces (O'Toole et al., 2002; O'Toole et al., 2006). Because both static and dynamic face recognition studies presented stimuli for 5 seconds each, participants would have considerably less time to process more visual information in the dynamic face recognition task. Therefore, even if dynamic information had a positive impact on participants' recognition performance, the advantage for recognizing dynamic identities may have been diminished by the fact that the recognition of unfamiliar dynamic faces may be

more challenging than the recognition of posed images identical with those presented in the learning phase.

Additionally, dynamic information appears to confer more advantage for recognizing other-race than own-race faces. The other-race effect may have been reduced because dynamic stimuli offer additional information, such as dynamic facial signatures or head movements, which participants can rely on to achieve recognition in addition to the invariant structures of the face. Cross-cultural studies have found that people of different races use different facial features to describe (Ellis, Deregowski, & Shepherd, 1975) and recognize (Blais et al., 2008) faces; furthermore, training observers to focus on the critical features relevant to a specific race of face (e.g. eyes for Caucasians and lips for Africans) may reduce the other-race effect, suggesting that observers may have been using the wrong looking strategy to perceive other-race faces, resulting in poorer recognition performance (Hills & Lewis, 2006). Previous static face recognition study has found that a combination of Eastern and Western eye movement strategies enabled Malaysian Chinese participants to recognize East Asian and Western Caucasian faces equally well, but not less familiar African faces (Tan et al., 2012). Because static images offer limited visual information, participants have to rely on the invariant structure of the face to achieve recognition; therefore, focusing on the diagnostic features of the face would have direct impact on participants' recognition performance. On the other hand, dynamic stimuli offer ample visual cues that participants can rely on to achieve recognition. Besides focusing on the diagnostic features relevant to a specific race of face, participants may recognize identities based on other information, such as one's characteristic crooked smile or head tilt, resulting in better recognition performance. This

indicates that attending to the lower regions of the face when perceiving dynamic stimuli may have enhanced Malaysian Chinese participants' recognition of the less familiar African faces as previous studies have identified the nose and mouth as critical features of African faces that, when attended, would reduce the other-race effect (Ellis et al., 1975; Hill & Lewis, 2006); additionally, dynamic stimuli may, perhaps, reduce the other-race effect by providing additional diagnostic information that is useful for recognizing identities, especially those from groups who are encountered less frequently.

In Chapter Six, results revealed that participants were able to identify all emotions described by models at above chance level. Recognizing emotions from muted videos of models describing an emotional experience may be more a more difficult task than recognizing posed images that capture a single portrayal of an intense emotion as models' display of emotions may be governed by different display rules relevant to their respective cultures; for example, models may mask their fear with a nervous laugh, de-intensify their sadness when recalling an emotional experience, or over-intensify their laughter to show happiness (Ekman, 1970). Even so, Malaysian Chinese participants were able to recognize all emotions, at above chance level, possibly because dynamic stimuli offered observers a perception of change in expressions, a factor which has been found to enhance emotion recognition in both static and dynamic stimuli (Ambadar et al., 2005).

7.4. Limitations and Suggestions for Future Research

This thesis only focused on the behavioral and eye tracking performance of Malaysian Chinese adults. The current findings were compared with the results of

previous eye tracking studies with similar aims and procedures (e.g. Blais et al., 2008; Jack et al., 2009); however, including participants of different races (e.g. Western Caucasians or Chinese participants with minimal exposure to Western culture) as comparison groups would allow direct comparisons of the results. Additionally, the current study only included demographic questionnaires which provided information about participants' duration of stay in different countries (i.e. participants for all four studies have not lived outside of their country for more than 3 years). Including more detailed questionnaires to identify the quality and quantity of participants' interracial contact and exposure to the West would provide more in-depth understanding of participants' contact level with people from different racial backgrounds and participants' familiarity with different cultures.

Initially, the static face recognition and emotion recognition task (presented in Chapter Three and Chapter Four) aimed to investigate Malaysian Chinese participants' recognition abilities and eye movement strategies for own- and other-race faces (i.e. East Asian and Western Caucasian faces only, a procedure similar to those employed in Blais et al. (2008) and Jack et al. (2009)'s studies). For the static emotion recognition task, a set of East Asian and Western Caucasian images were collected by recruiting students and staffs from the University of Nottingham Malaysia Campus. As for the static face recognition task, African stimuli collected from the University of Pretoria, South Africa were included in addition to the East Asian and Western Caucasian stimuli obtained from a set of images collected at the University of St Andrews, UK. Because the behavioral and eye tracking results for the static face recognition task revealed differences in the way participants perceive and recognize other-race faces based on familiarity with the

culture, the remaining dynamic face recognition and emotion recognition tasks included African stimuli in addition to East Asian and Western Caucasian ones to explore these differences. Unfortunately, participants' behavioral and eye tracking performances for African stimuli were not explored in the static emotion recognition task.

For both dynamic face recognition and emotion recognition tasks, the age of Western Caucasian models was older than East Asian models. Because Western Caucasians (and Africans) consist of a smaller proportion of the student population, most models were recruited by opportunity sampling and snowball sampling, which resulted in more staff volunteers for Western Caucasian models. However, all models were under the age of 45. Brain-imaging studies examining the behavioral and neural representation of facial expressions across different age groups classified participants aged 50 and above as older adults, suggesting that the behavioral and neural responses in adults under the cut off age should be similar (Somerwille, Fani, & McClure-Tone, 2012). Therefore, although Western Caucasian models were older than East Asians, models should express their emotions and present themselves in a similar manner, which would not affect participants' performance when perceiving these faces. However, future studies should aim to recruit models of similar age to ensure that models' age does not affect participants' performance.

Future research could focus more on the development of naturalistic stimuli. The dynamic emotion recognition task presented in Chapter Six aimed to create naturalistic stimuli which more accurately represent a typical social situation by presenting muted videos of models describing emotional experiences. However, future studies should further investigate whether models were actually expressing the emotions being described

and whether participants rely on these expressions to achieve recognition. In the current study, models were instructed to describe an experience related to an assigned emotion, followed by self-ratings of the intensity of emotion felt during the narrated experience and the descriptiveness of the narration. Analyses revealed that the intensity of emotion and descriptiveness of narration did not differ between all emotions and models' racial background, suggesting no clear advantage for participants to recognize certain emotions more accurately than others. Presumably, models have expressed diagnostic information related to the emotion described, which allowed participants to correctly identify all emotions above chance level based on visual information alone. However, to address this question directly, future studies could, perhaps, use a standardized system (e.g. the facial action coding system) to categorize the emotions expressed by classifying each facial movement individually. Moreover, further research should investigate whether different presentation methods, such as posed expressions versus natural introduction and description of emotions that more accurately represent a typical social interaction, would affect participants' behavioral and eye tracking performance.

Future research should also examine whether the individualist-collectivist nature of a culture, the exposure to and familiarity with faces, or aspects of both affect observers' eye movement strategy when perceiving faces. Although it is possible that the Eastern and Western eye movement strategies identified in face perception may be an extension of the holistic-analytic cognitive and perceptual style found in the respective cultures, studies exploring the effect of environment have revealed that Asian participants who are more familiar with Western culture adopted a unique fixation pattern different from those observed in East Asians, which suggests that there may be more to cross-cultural

differences than a simple Eastern-Western distinction. Research suggests that exposure to and familiarity with the diagnostic features of other-race faces encountered during development may also affect one's looking pattern. Future studies exploring the relationship between participants' perceptual style, exposure to other-race faces, and eye movement strategy will provide insight to the cross-cultural differences in eye movement strategies employed in face processing tasks.

7.5. Conclusion

In summary, the current thesis revealed that possibly due to greater exposure to Western culture, Malaysian Chinese adults adopted a combination of both Eastern and Western eye movement strategies (with more fixations on the eyes, followed by the nose then mouth) to better recognize other-race identities and emotions, implying that directing attention towards facial cues that have the highest diagnostic value in relation to a specific race of face and/or task may subsequently reduce the other-race effect and improve participants' recognition performance. Additionally, the findings revealed that the type of stimuli (i.e. static versus dynamic) and nature of the task (i.e. face recognition versus emotion recognition) rendered different eye movement strategies from Malaysian Chinese participants, providing further support that participants perform differently when presented with dynamic stimuli that are more naturalistic and when instructed to perform different tasks. Because research has found that studies utilizing static stimuli lack ecological validity, the dynamic face recognition and emotion recognition tasks aimed to address various methodological concerns by developing naturalistic stimuli that more accurately represent a typical social situation. Therefore, the behavioral and eye tracking

performance observed in the dynamic face recognition and emotion recognition tasks may be a more accurate reflection of observers' recognition abilities and looking strategies in real life settings; more specifically, observers may show a preference for the lower regions of the face (i.e. the nose and mouth area) rather than the eyes when perceiving faces.

References

Ackerman, J. M., Shapiro, J. R., Neuberg, S. L., Kenrick, D. T., Vaughn Becker, D., Griskevicius, V., Maner, J. K., & Schaller, M. They all look the same to me (unless they're angry). *Psychological Science*, *17 (10)*, 836-840.

Advameg (2011). Malaysia. *Countries and Their Cultures*. Retrieved October 10, 2011, from www.everyculture.com/Ja-Ma/Malaysia.html

Ambadar, Z., Schooler, J. W., & Cohn, J. F. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science*, *16 (5)*, 403-410.

Andrews, T. J. & Coppola, D. M. (1999). Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision Research*, *39*, 2947-2953.

Background Note: Malaysia. *U.S. Department of State*. (2011). Retrieved October 10, 2011, from www.state.gov/r/pa/ei/bgn/2777.htm

Biehl, M., Matsumoto, D., Ekman, P., Hearn, V., Heider, K., Kudoh, T., & Ton, V. (1997). Matsumoto and Ekman's Japanese and Caucasian Facial Expressions of Emotion (JACFEE): Reliability data and cross-national differences. *Journal of Nonverbal Behavior*, *21*, 3-21. doi:10.1023/A:1024902500935

Birdwhistell, R. (1970). Kinesics and Context. Philadelphia, PA: University of Pennsylvania Press.

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. PLoS ONE 3 (8): e3022. doi:10.1371/journal.pone.0003022.

Blais, C., Roy, C., Fiset, D., Arguin, M., & Gosselin, F. (2012). The eyes are not the window to basic emotions. *Neuropsychologia*, *50*, 2830-2838.

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

Bradshaw, J. L. & Wallace, G. (1971). Models for the processing and identification of faces. *Perception and Psychophysics*, *9 (5)*, 443-448.

Bruce, V. & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77 (3)*, 305-327.

Bruyer, R. (1991). Covert face recognition in prosopagnosia: A review. *Brain and Cognition*, *15*, 223-235.

Bruyer, R. (2011). Configural face processing: A meta-analytic survey. *Perception*, *40*, 1478-1490.

Buchan, J. N., Pare, M., & Munhall, K. G. (2007). Spatial statistics of gaze fixations during dynamic face processing. *Social Neuroscience, 2 (1)*, 1-13.

Burton, A. M., Bruce, V., & Johnston, R. A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology*, *81*, 361-380.

Burton, A. M., Young, A. W., Bruce, V., Johnston, R. A., & Ellis, A. W. (1991). Understanding covert recognition. *Cognition*, *39*, 129-166.

Caldara, R., Zhou, X., & Miellet, S. (2010). Putting culture under the 'spotlight' reveals universal use for face recognition. PLoS ONE 5(3): e9708. doi:10.1371/journal.pone.0009708

Choi, I. & Nisbett, R. E. (1998). Situational salience and cultural differences in the correspondence bias and actor-observer bias. *Personality and Social Psychology Bulletin, 24 (9),* 949-960.

Chua, H. F., Boland, J. E., & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *PNAS, 102 (35),* 12629-12633.

Davies, A. G. (2011). Radio stations in Malaysia. *Asiawaves*. Retrieved August 25, 2011, from http://www.asiawaves.net/malaysia-radio.htm

Davies, G., Ellis, H., & Shepherd, J. (1977). Cue saliency in faces as assessed by the 'Photofit' technique. *Perception, 6 (3),* 263-269.

Darwin, C. (1872). *The Expression of the Emotions is Man and Animals*. New York: Philosophical Library.

De Haan, E. H. F., Young, A., & Newcombe, F. (1987). Faces interfere with name classification in a prosopagnosic patient. *Cortex, 23,* 309-316.

Doherty, M., Tsuji, H., & Phillips, W. A. (2008). The context sensitivity of visual size perception varies across cultures. *Perception, 37,* 1426-1433.

Ekman, P. (1970). Universal facial expressions of emotion. *California Mental Health Research Digest, 8 (4),* 151-158.

Ekman, P. & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17 (2),* 124-129.

Ellis, H. D., Deregowski, J. B., & Shepherd, J. W. (1975). Descriptions of White and Black faces by White and Black subjects. *International Journal of Psychology, 10*, 119-123.

Ellis, H. D., Sheperd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception, 8*, 431-439.

Epstein, J. (2011).World Domination by Box Office Cinema Admissions. Greenash. Retrieved August 25, 2011, from http://greenash.net.au/thoughts/2011/07/world-domination-by-box-office-cinema-admissions/.

Ethnicity and Race by Countries. (2005). Retrieved February 10, 2014, from http://www.infoplease.com/ipa/A0855617.html.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review*, *105 (3)*, 482-498.

Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants, *Child Development, 64*, 637-656.

Fiorentini, C. & Viviani, P. (2011). Is there a dynamic advantage for facial expressions? *Journal of Vision, 11 (3): 17*, 1-15.

Friesen, W. V., 1972. Cultural differences in facial expression in a social situation: An experimental test of the concept of display rules. Unpublished doctoral dissertation. University of California, San Francisco.

Gaudart, H. (1987). English language teaching in Malaysia: A historical account. *The English Teacher* Vol XVI.

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin, 75 (6),* 424-429.

Guo, K. (2012). Holistic gaze strategy to categorize facial expression of varying intensities. PLoS ONE 7(8): e42585. doi:10.1371/journal.pone.0042585.

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

Happé, F. & Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 36 (1),* 5-25.

Hayhoe, M. & Ballard, D. (2005). Eye movements in natural behavior. *TRENDS in Cognitive Sciences, 9 (4),* 188-194.

Heine, S. J., Lehman, D. R., Markus, H. R., & Kitayama, S. (1999). Is there a universal need for positive self-regard. *Psychological Review, 106,* 766-794.

Heron-Delaney, M., Anzures, G., Herbert, J. S., Quinn, P. C., Slater, A. M., Tanaka, J. W., Lee, K., & Pascalis, O. (2011). Perceptual training prevents the emergence of the other race effect during infancy. PLoS ONE 6 (5): e19858. doi:10.1371/journal.pone.0019858.

Hess, U. & Kleck, R. E. (1990). Differentiating emotion elicited and deliberate emotional facial expressions. *European Journal of Social Psychology*, *20,* 369-385.

Hill, H. & Johnston, A. (2011). Categorizing sex and identity from the biological motion of faces. *Current Biology*, *11*, 880-885.

Hills, P. J. & Lewis, M. B. (2006). Reducing the own-race bias in face recognition by shifting attention. *The Quarterly Journal of Experimental Psychology, 59 (6),* 996-1002.

Hill, H., Schyns, P. G., & Akamatsu, S. (1997). Information and viewpoint dependence in face recognition. *Cognition*, *62*, 201-222.

Hockley, W. E. Hemsworth, D. H., Consoli, A. (1999). Shades of the mirror effect: Recognition of faces with and without sunglasses. *Memory and Cognition, 27 (1),* 128-138.

Hoffmann, H., Traue, H. C., Limbrecht-Ecklundt, K., Walter, S., & Kessler, H. (2013). Static and dynamic presentation of emotions in different facial areas: Fear and surprise show influences of temporal and spatial properties. *Scientific Research*, *4 (8)*, 663-668.

Horstmann, G. & Ansorge, U. (2009). Visual search for facial expressions of emotions: A comparison of dynamic and static faces. *Emotion, 9 (1),* 29-38.

Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology, 19,* 1543-1548.

Jack, R. E., Caldara, R., & Schyns, P. G. (2011). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology*, *11*, 1-7. doi:10.1037/a0023463

Ji, L.J., Peng, K., & Nisbett, R. E. (2000). Culture, control, and perception of relationships in environment. *Journal of Personality and Social Psychology*, *78*, 943-955.

Jones, E. E., & Harris, V. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology, 3,* 1-24.

Kanapathy, V. (2008). Malaysia. Asian and Pacific Migration J 17(3–4): 335–347.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2,* 217-250.

Kelly, D. J., Jack, R. E., Miellet, S., De Luca, E., Foreman, K., & Caldara, R. (2011a). Social experience does not abolish cultural diversity in eye movements. *Frontiers in Psychology, 2 (95),* 1-11.

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*, 105-114.

Kelly, D. J., Liu, S., Rodger, H., Miellet, S., Ge, L., & Caldara, R. (2011b). Developing cultural differences in face processing. *Developmental Science*, *14 (5)*, 1176-1184.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy. *Psychological Science, 18 (12),* 1084-1089.

Kessler, H., Doyen-Waldecker, C., Hofer, C., Hoffmann, H., Traue, H. C., & Abler, B. (2011). Neural correlates of the perception of dynamic versus static facial expressions of emotion. *GMS Psycho-Social Medicine, 8,* 1-8.

Kleinschmidt, A. & Cohen, L. (2006). The neural bases of prosopagnosia and pure alexia: Recent insights from functional neuroimaging. *Current Opinion in Neurology, 19,* 386-391.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Arch Gen Psychiatry, 59,* 809-816.

Klineberg, O. (1940). *Social Psychology*. New York: Holt.

Knappmeyer, B., Thornton, I. M., & Bulthoff, H. H. (2003). The use of facial motion and facial form during the processing of identity. *Vision Research, 43,* 1921-1936.

Knight, B. & Johnston, A. (1997). The role of movement in face recognition. *Visual Cognition, 4 (3),* 265-273.

Lander, K. & Bruce, V. (2000). Recognizing famous faces: Exploring the benefits of facial motion. *Ecological Psychology, 12 (4),* 259-272.

Lander, K., Christie, F., & Bruce, V. (1999). The role of movement in the recognition of famous faces. *Memory and Cognition, 27 (6),* 974-985.

Liu, S., Quinn, P. C., Wheeler, A., Xiao, N., Ge, L., & Lee, K. (2011). Similarity and difference in the processing of same- and other-race faces as revealed by eye

tracking in 4- to 9-month-olds. *Journal of Experimental Child Psychology*, *108*, 180-189.

Masuda, T. & Nisbett, R.E. (2001). Attending holistically versus analytically: comparing the context sensitivity of Japanese and Americans. *Journal of Personality and Social Psychology*, *81*, 922-934.

Matsumoto, D. & Willingham, B. (2009). Spontaneous facial expressions of emotion of congenitally and noncongenitally blind individuals. *Journal of Personality and Social Psychology*, *96*, 1-10. doi:10.1037/a0014037

Mealey, L., Daood, C., & Krage, M. (1996). Enhanced memory for faces of cheaters. *Ethology and Sociobiology*, *17*, 119-128.

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

Miyamoto, Y., Nisbett, R. E., & Masuda, T. (2006). Culture and the physical environment. Holistic versus analytic perceptual affordances. *Psychological Science, 17*, 113-119.

Moriguchi, Y., Ohnishi, T., Kawachi, T., Mori, T., Hirakata, M., Yamada, M., Matsuda, H., & Komaki, G. (2005). Specific brain activation in Japanese and Caucasian people to fearful faces. *NeuroReport*, *16*, 133-136.

Newcombe, F., Young, A. W., & De Haan, E. H. F. (1989). Prosopagnosia and object agnosia without cover recognition. *Neuropsychologia, 27 (2)*, 179-191.

Nisbett, R. E. & Miyamoto, Y. (2005). The influence of culture: holistic versus analytic perception. *TRENDS in Cognitive Sciences, 9 (10),* 467-473.

Norenzayan, A., Choi, I., & Nisbett, R. E. (2002). Cultural similarities and differences in social inference: Evidence from behavioral predictions and lay theories of behavior. *Personality and Social Psychology Bulletin, 28 (1),* 109-120.

O'Toole, A. J., Jiang, F., Roark, D., & Abdi, H. (2006). Predicting human performance for face recognition. In Zhao, W., CHellapppa, R. (Eds.), Face processing: Advanced modelling and methods (293-319). New York: Academic Press.

O'Toole, A. J., Roark, D. A., & Abdi, H. (2002). Recognizing moving faces: A psychological and neural synthesis. *TRENDS in Cognitive Sciences, 6 (6),* 261-266.

Pallier, C. (2002). Computing discriminability and bias with the R software. Retrieved August 25, 2011, from www.pallier.org

Papesh, M. H. & Goldinger, S. D. (2009). Deficites in other-race face recognition: No evidence for encoding-based effects. *Canadian Journal of Experimental Psychology, 63 (4),* 253-262. doi:10.1037/a0015802.

Peterson, M. A. & Rhodes, G. (2003). Analytic and holistic processing: The view through different lenses. In M.A. Peterson & G. Rhodes (Eds.), *Perception of Faces, Objects, and Scenes: Analytic and Holistic Processes*. Pp. 3-19. New York: Oxford University Press.

Pezdek, K. (2006). Expert testimony on eyewitness memory and identification. Costanzo, M., Krauss, D., & Pezdek, K. (Eds.), *Expert psychological testimony for the courts* (pp.99-117). Mahwah, NJ: Erlbaum.

Piepers, D. W. & Robbins R. A. (2012). A review and clarification of the terms "holistic," "configural," and "relational" in the face perception literature. *Frontiers in Psychology, 3: 559*.

Pillai, D., Sheppard, E., & Mitchell, P. (2012). Can people guess what happened to others from their reactions? PLoS ONE 7 (11): e49859. doi:10.1371/journal.pone.0049859.

Pozzulo, J. D. & Lindsay, R. C. L. (1998). Identification accuracy of children versus adults: a meta-analysis. *Law and Human Behavior, 22*, 549-57.

Read, J. D., Vokey, J. R., & Hammersley, R. (1990). Changing photos of faces: Effects of exposure duration and photo similarity on recognition and the accuracy-confidence relationship. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16 (5)*, 870-882.

Rhodes, G. (1988). Looking at faces: First-order and second-order features as determinates of facial appearance. *Perception, 17*, 43-63.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin, 115 (10)*, 102-141.

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.

Senju, A., Vernetti, A., Kikuchi, Y., Akechi, H., & Hasegawa, T. (2013a). Cultural modulation of face and gaze scanning in young children. PLoS ONE 8 (8): e74017. doi:10.1371/journal.pone.0074017.

Senju, A., Vernetti, A., Kikuchi, Y., Akechi, H., Hasegawa, T., & Johnson, M. H. (2013b). Cultural background modulates how we look at other persons' gaze. *International Journal of Behavioral Development, 37,* 13-136.

Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology, 75*, 221-242.

Somerville, L. H., Fani, N., & McClure-Tone, E. B. (2012). Behavioral and neural representation of emotional facial expressions across the lifespan. *Developmental Neuropsychology, 36 (4),* 408-428.

Stanley, J. T., Zhang, X., Fung, H. H., & Isaacowitz, D. M. (2013). Cultural differences in gaze and emotion recognition: Americans contrast more than Chinese. *Emotion, 13 (1),* 36-46.

Stephen, I. D., Coetzee, V., Law Smith, M., & Perrett, D. I. (2009). Skin blood perfusion and oxygenation colour affect perceived human health. PLoS ONE 4 (4): e5083. doi:10.1371/journal.pone.0005083.

Tan, C. B. Y., Stephen, I. D., Whitehead, R., & Shepard, E. (2012). You look familiar: How Malaysian Chinese recognize faces. PLoS ONE 7 (1):e29714. doi:10.1371/journal.pone.0029713.

Tanaka, J. W. & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology, 46A (2),* 225-245.

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

Tardif, T., Gelman, S. A., & Xu, F. (1999). Putting the "noun bias" in context: A comparison of English and Mandarin. *Child Development, 70 (3),* 620-635.

The World Factbook. *Central Intelligence Agency*. (2011). Retrieved August 25, 2011, from https://www.cia.gov/library/publications/the-world-factbook/geos/my.html

Tiddeman, B. & Perrett, D. (2001). Moving Facial Image Transformations Based On Static 2D Prototypes, *Proc. 9th Int. Conf. In Central Europe on Computer Graphics, Visualization and Computer Vision 2001,* Pilsen, Czech Republic, Feb 5-9, 2001.

Turati, C., Di Giorgio, E., Bardi, L., & Simion, F. (2010). Holistic face processing in newborns, 3-month-old infants, and adults: Evidence from the composite face effect. *Child Development, 81 (6),* 1894-1905.

Valentin, D., Abdi, H., & Edelman, B. (1999). From rotation to disfiguration: Testing a dual-strategy model for recognition of faces across view angles. *Perception, 28,* 1-10.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology, 43A (2),* 161-204.

Vo, M. L. H., Smith, T. J., Mital, P.K., & Henderson, J. M. (2012). Do the eyes really have it? Dynamic allocation of attention when viewing moving faces. *Journal of Vision, 12 (13): 3,* 1-14.

Volstorf, J. Rieskamp, J., & Stevens, J. R. (2011). The good, the bad, and the rare: Memory for partners in social interactions. PLoS ONE 6 (4): e18945. doi:10.1371/journal.pone.0018945.

Wheeler, A., Anzures, G., Quinn, P. C., Pascalis, O., Omrin, D. S., & Lee, K. (2011). Caucasian infants scan own- and other-race faces differently. *PLoS ONE, 6 (4),* 1-8. doi:10.1371/journal.pone.0018621

Wehrle, T., Kaiser, S., Schmidt, S., & Scherer, K. R. (2000). Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality and Social Psychology, 78 (1),* 105-119.

Wells, G. L. & Olson, E. A. (2003). Eyewitness testimony. *Annual Reviews Psychology, 54,* 277-295.

Wickens, A. P. (2009). *Introduction to Biopsychology*. Harlow, Essex: Pearson Education Limited.

Yarbus, A. (1967). Eye movements and vision. Plenum Press: New York.

Young, A., Hellawell, D., & Hay, D. (1987). Configurational information in face perception. *Perception, 16,* 747-759.

Young, A. W. & De Haan, E. H. F. (1988). Boundaries of covert recognition in prosopagnosia. *Cognitive Neuropsychology, 5,* 317-336.

Yuki, M., Maddux, W. W., & Masuda, T. (2007). Are the windows to the soul the same in the East and West? Cultural differences in using the eyes and mouth as cues to recognize emotions in Japan and the United States. *Journal of Experimental Social Psychology, 43,* 303-311.

Zhang, J. & Mueller, S. T. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika, 70 (1),* 145-154.

Appendix A

Results for Total Fixation Duration in the Static Face Recognition Task

For total fixation duration, there was a main effect of race, $F(2, 5343.34)=5.25$, $p=.005$, feature, $F(2, 5343.34)=736.71$, $p<.001$, and interactions between task and feature, $F(2, 5343.34)=5.11$, $p=.006$, and race and feature, $F(4, 5343.34)=4.80$, $p=.001$, see Figure A.1. For race, post-hoc tests using Bonferroni Pairwise Comparisons showed that the total fixation duration spent on African faces was significantly longer than Western Caucasian ($p=.024$) and East Asian ($p=.01$) faces. However, the total fixation duration spent on Western Caucasian and East Asian faces did not differ significantly ($p=1$). As for feature, post-hoc tests using Bonferroni Pairwise Comparisons showed that the total fixation duration landing on the eyes was significantly longer than the nose ($p=.001$) and mouth ($p<.001$), and the nose longer than the mouth ($p<.001$).

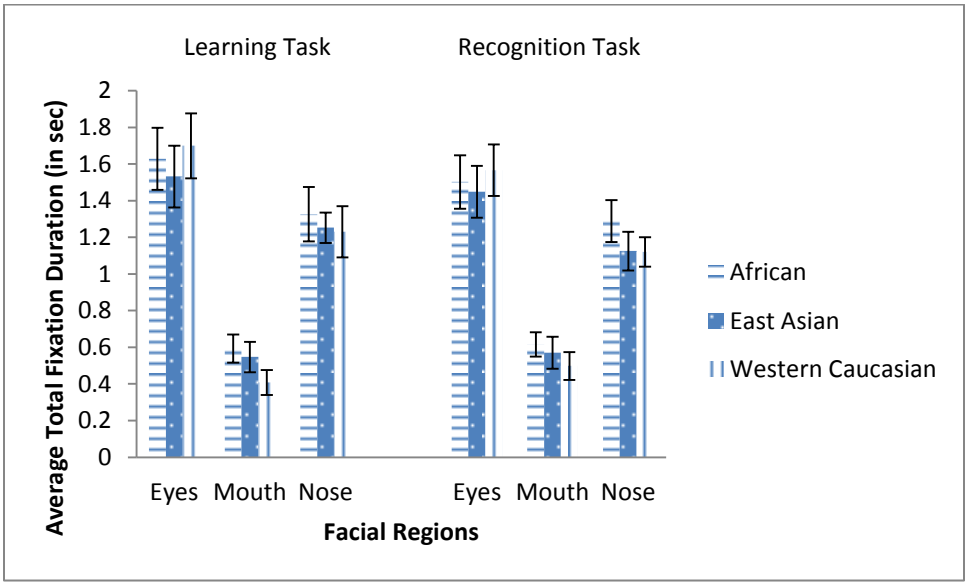


Figure A.1. Average total fixation duration for the eyes, mouth, and nose during the learning and recognition tasks, in seconds. Error bars report standard errors of mean. The total fixation duration landing on the eyes was longer than the nose and mouth, and the nose longer than the mouth.

To examine the interaction between task and feature, separate Linear Mixed Models were conducted for each of the two tasks. Participant ID, face ID, and feature were included as factors (dependent variable = total fixation duration; fixed factor = feature of face). There was a main effect of feature for the learning task, $F(2, 1797)=278.83$, $p<.001$, and for the recognition, $F(2, 3597)=453.41$, $p<.001$, task. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for both tasks, the total fixation duration spent looking at the eyes was longer than the nose ($p<.001$ for both tasks) and mouth ($p<.001$ for both tasks), and the nose longer than the mouth ($p<.001$ for both tasks). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and task were included as factors (dependent variable = total fixation duration; fixed factor = task). There was a main effect of task for eyes, $F(1, 1798)=5.64$, $p=.018$, and nose, $F(1, 1798)=5.19$, $p=.023$, but not for the mouth, $F(1, 1798)=2.66$, $p=.103$. Post-hoc tests using Bonferroni Pairwise Comparisons showed that the total fixation duration landing on the eyes ($p=.018$) and nose ($p=.023$) were significantly longer for the learning than recognition task.

To examine the interaction between race and feature, separate Linear Mixed Models were conducted for each of the three races. Participant ID, face ID, and feature were included as factors (dependent variable = total fixation duration; fixed factor = feature of face). There was a main effect of feature for African, $F(2, 1797)=207.89$,

$p<.001$, East Asian, $F(2, 1797)=212.91$, $p<.001$, and Western Caucasian faces, $F(2, 1797)=327.12$, $p<.001$. Post-hoc tests using Bonferroni Pairwise Comparisons showed that for all three races of faces, the total fixation duration spent looking at the eyes was significantly longer than the nose ($p<.001$ for all three races) and mouth ($p<.001$ for all three races), and the nose longer than the mouth ($p<.001$ for all three races). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and race were included as factors (dependent variable = total fixation duration; fixed factor = race of face). There was a main effect of race for mouth, $F(2, 1797)=9.80$, $p<.001$, and nose, $F(2, 1797)=6.02$, $p=.002$, and a trend for the eyes, $F(2, 1797)=2.91$, $p=.055$. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for the mouth, the total fixation duration was significantly longer for African ($p<.001$) and East Asian ($p=.01$) than Western Caucasian faces. However, the total fixation duration on the mouth did not differ significantly for African and East Asian faces ($p=.473$). For the nose, the total fixation duration was significantly longer for African than East Asian ($p=.012$) and Western Caucasian faces ($p=.006$). However, the total fixation duration on the nose did not differ significantly for East Asian and Western Caucasian faces ($p=1$). As for the eyes, the total fixation duration was significantly longer for Western Caucasian than East Asian faces ($p=.048$). However, the total fixation duration on the eyes for African faces did not differ from Western Caucasian ($p=.702$) and East Asian faces ($p=.665$).

Appendix B

Results for Total Fixation Duration in the Static Emotion Recognition Task

For total fixation duration, there was a main effect of emotion, $F(6, 4158)=2.82$, $p=.01$, feature, $F(2, 4158)=380.95$, $p<.001$, an interaction between emotion and feature, $F(12, 4158)=5.20$, $p<.001$, and an interaction between race, emotion, and feature, $F(12, 4158)=2.01$, $p=.02$, see Figure B.1. For emotion, post-hoc tests using Bonferroni Pairwise Comparisons revealed that the total fixation duration for "fearful" faces was significantly longer than "happy" faces ($p=.003$), and a trend towards longer total duration on "sad" than "happy" ($p=.071$) faces. As for feature, post-hoc tests using Bonferroni Pairwise Comparisons showed that the total fixation duration on the eyes was significantly longer than the nose ($p<.001$) and mouth ($p<.001$), and the nose longer than the mouth ($p<.001$).

To examine the interaction between emotion and feature, separate Linear Mixed Models were conducted for each of the seven emotions. Participant ID, face ID, and feature were included as factors (dependent variable = total fixation duration; fixed factor = feature of face). There was a main effect of feature for all seven emotions, $F(2, 597)=81.75$, $p<.001$ for "neutral," $F(2, 597)=29.65$, $p<.001$ for "happy," $F(2, 597)=29.21$, $p<.001$ for "surprise," $F(2, 597)=89.88$, $p<.001$ for "fear," $F(2, 597)=51.15$, $p<.001$ for "disgust," $F(2, 597)=65.81$, $p<.001$ for "anger," and $F(2, 597)=64.94$, $p<.001$ for "sad" faces. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for "neutral," "happy," "surprise," "fear," and "anger" faces, the total fixation duration on the eyes was significantly longer than the nose ($p<.001$ for "neutral" and "fear" faces, $p=.041$ for "happy" faces, $p=.003$ for "surprise"

faces, and p=.038 for "anger" faces) and mouth (p<.001 for all five emotions), and the nose longer than the mouth (p<.001 for all five emotions). For "disgust" and "sad" faces, the total fixation duration on the eyes (p<.001 for both emotions) and nose (p<.001 for both emotions) was significantly longer than the mouth; however, the total fixation duration on the eyes and nose did not differ significantly (p=.543 for "disgust" faces and p=.982 for "sad" faces). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and emotion were included as factors (dependent variable = total fixation duration; fixed factor = emotion). There was a main effect of emotion for the eyes, F(6, 1393)=5.49, p<.001, and nose, F(6, 1393)=4.35, p<.001, and a trend towards a main effect for the mouth, F(6, 1393)=2.09, p=.052. Post-hoc tests using Bonferroni Pairwise Comparisons showed that the total fixation duration on the eyes was significantly longer for "fear" than "happy" (p<.001), "surprise" (p=.019) and "disgust" (p=.001) faces, and a longer total fixation duration for "neutral" than "happy" (p=.006) and "disgust" faces (p=.012). For the nose, the total fixation duration on the nose was significantly longer for "disgust" than "happy" (p=.002) and "surprise" (p=.008) faces, and "sad" than "happy" (p=.011), and "surprise" (p=.038) faces. As for the mouth, there was a trend towards longer total fixation duration on "surprise" than "neutral" (p=.053) and "anger" (p=.074) faces.

To examine the interaction between race, emotion, and feature, two separate Linear Mixed Models were conducted for each race. For East Asian faces, there was a main effect of feature, F(2, 2079)=177.95, p<.001, a trend towards a main effect of emotion, F(6, 2079)=1.97, p=.067, and an interaction between emotion and feature, F(12, 2079)=4.19, p<.001. For feature, post-hoc tests using Bonferroni Pairwise Comparisons revealed that the total fixation duration on East Asian eyes was

significantly longer than the nose (p<.001) and mouth (p<.001), and the nose significantly longer than the mouth (p<.001). However, Bonferroni Pairwise Comparisons failed to reach significance for all emotions. To examine the interaction between emotion and feature for East Asian faces, separate Linear Mixed Models were conducted for each of the seven emotions. Participant ID, face ID, and feature were included as factors (dependent variable = total fixation duration; fixed factor = feature of face). There was a main effect of feature for all seven emotions, F(2, 297)=47.582, p<.001 for "neutral," F(2, 297)=10.64, p<.001 for "happy," F(2, 297)=9.74, p<.001 for "surprise," F(2, 297)=39.46, p<.001 for "fear," F(2, 297)=21.91, p<.001 for "disgust," F(2, 297)=46.22, p<.001 for "anger," and F(2, 297)=24.35, p<.001 for "sad" faces. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for "neutral" and "fear" East Asian faces, the total fixation duration on the eyes was significantly longer than the nose (p<.001 for both emotions) and mouth (p<.001 for both emotions), and the nose longer than the mouth (p<.001 for both emotions). For "surprise," "disgust," "anger," and "sad" faces, the total fixation duration on the eyes (p<.001 for all four emotions) and nose (p<.001 for "disgust," "anger," and "sad" and p=.049 for "surprise" faces) was significantly longer than the mouth; however, the total fixation duration on the eyes and nose did not differ significantly (p=.143 for "surprise," p=1 for "disgust," "anger," and "sad" faces). For "happy" faces, the total fixation duration landing on the eyes was significantly longer than the nose (p=.011) and mouth (p<.001); however, the total fixation duration on the nose and mouth did not differ significantly (p=.313). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and emotion were included as factors (dependent variable = total fixation duration; fixed factor = emotion). There was a main effect of emotion for the eyes,

F(6, 693)=3.34, *p*=.003, mouth, *F*(6, 693)=3.41, *p*=.003, and nose, *F*(6, 693)=3.61, *p*=.002. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for the eyes, the total fixation duration landing on "fear" faces was significantly longer than "happy" (*p*=.049) and "surprise" (*p*=.039) faces, "neutral" longer than "surprise" (*p*=.049) faces, and a trend towards a longer total fixation duration on "neutral" than "happy" faces (*p*=.06). For the mouth, the total fixation duration landing on "happy" (*p*=.007) and "surprise" (*p*=.006) faces was significantly longer than "angry" faces. As for the nose, the total fixation duration landing on "disgust" (p=.027), "anger" (*p*=.006), and "sad" (*p*=.038) faces was significantly longer than "happy" faces

For Western Caucasian faces, there was a main effect of feature, *F*(2, 2079)=203.74, *p*<.001, and an interaction between emotion and feature, *F*(12, 2079)=3.03, *p*<.001. For feature, post-hoc tests using Bonferroni Pairwise Comparisons revealed that the total fixation duration on Western Caucasian eyes was significantly longer than the nose (*p*<.001) and mouth (*p*<.001), and the nose longer than the mouth (*p*<.001). To examine the interaction between emotion and feature for Western Caucasian faces, separate Linear Mixed Models were conducted for each of the seven emotions. Participant ID, face ID, and feature were included as factors (dependent variable = total fixation duration; fixed factor = feature of face). There was a main effect of feature for all seven emotions, *F*(2, 297)=34.72, *p*<.001 for "neutral," *F*(2, 297)=22.65, *p*<.001 for "happy," *F*(2, 297)=20.30, *p*<.001 for "surprise," *F*(2, 297)=51.04, *p*<.001 for "fear," *F*(2, 297)=30.71, *p*<.001 for "disgust," *F*(2, 297)=23.76, *p*<.001 for "anger," and *F*(2, 297)=41.58, *p*<.001 for "sad" faces. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for "neutral," "surprise," "fear" and "anger" Western Caucasian faces, the total fixation duration on the eyes was significantly longer than the nose (*p*=.002 for "neutral," *p*=.019 for

"surprise," $p<.001$ for "fear, " and $p=.014$ for "anger" faces) and mouth ($p<.001$ for all four emotions), and the nose longer than the mouth ($p<.001$ for all four emotions). For "happy," "disgust," and "sad" faces, the total fixation duration on the eyes ($p<.001$ for all three emotions) and nose ($p<.001$ for all three emotions) was significantly longer than the mouth; however, the total fixation duration on the eyes and nose did not differ significantly ($p=1$ for "happy" and "sad" faces and $p=.118$ for "disgust" faces). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and emotion were included as factors (dependent variable = total number of first fixations; fixed factor = emotion). There was a main effect of emotion for the eyes, $F(6, 693)=2.87$, $p=.009$ and nose, $F(6, 693)=2.41$, $p=.026$, but not for the mouth, $F(6, 693)=1.69$, $p=.121$. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for the eyes, the total fixation duration on "fear" faces was significantly longer than "happy" ($p=.039$) and "disgust" ($p=.012$) faces. As for the nose, there were no significant differences between all paired samples.

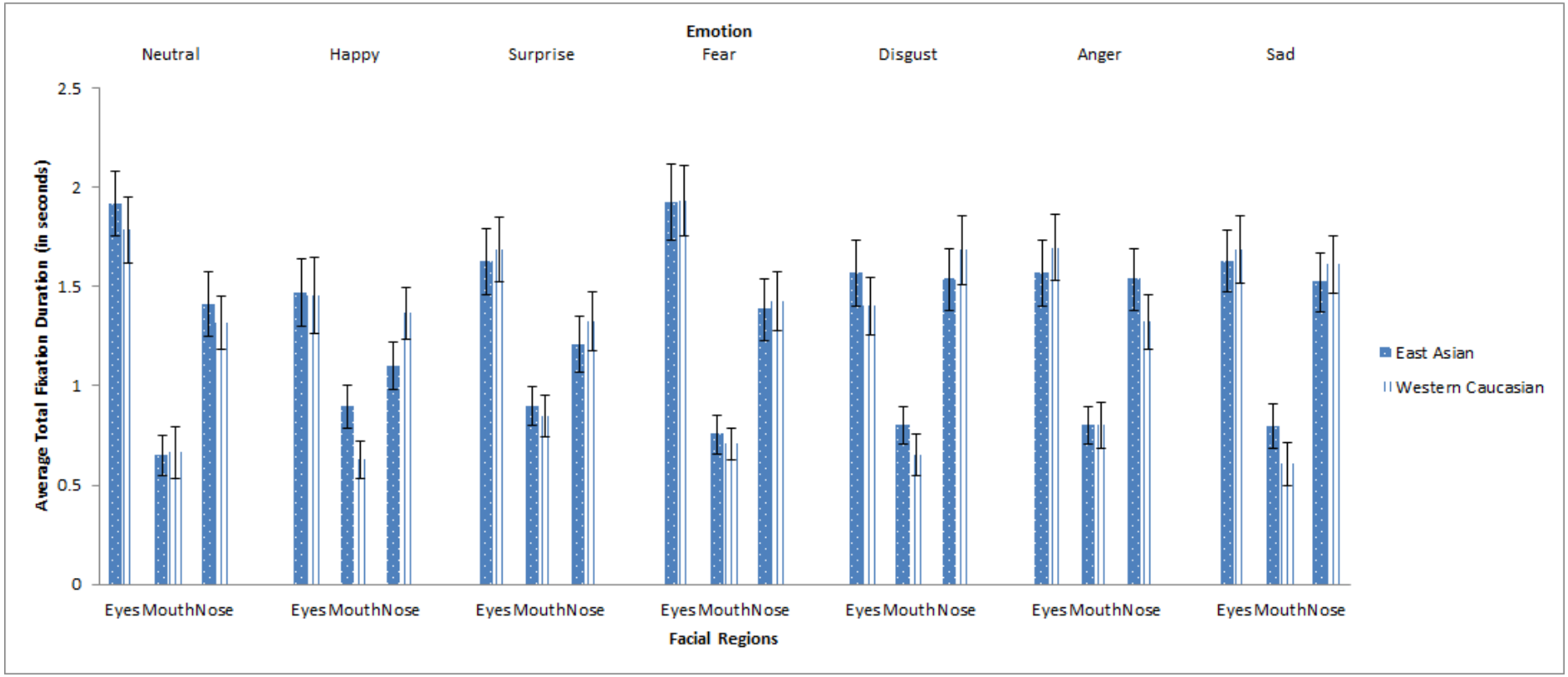


Figure B.1. Average total fixation duration for the eyes, mouth, and nose during the emotion recognition task, in seconds. Error bars report standard errors of mean. Participants fixated longer on the eyes than the nose and mouth, and the nose longer than the mouth.

Appendix C

Results for Total Fixation Duration in the Static Face Recognition and Emotion Recognition Tasks

For total fixation duration, there was a main effect of task, $F(1, 42.73)=22.50$, $p<.001$, feature, $F(2, 6543.81)=638.68$, $p<.001$, an interaction between task and feature, $F(2, 6543.81)=3.27$, $p=.038$ and a trend towards an interaction between race and feature, $F(2, 6543.84)=2.74$, $p=.064$, see Figure C.1. For task, the total fixation duration on the emotion recognition task was significantly longer than the face recognition task ($p<.001$). As for feature, post-hoc tests using Bonferroni Pairwise Comparisons showed that the total fixation duration on the eyes was significantly longer than the nose ($p<.001$) and mouth ($p<.001$), and the nose significantly longer than the mouth ($p<.001$).

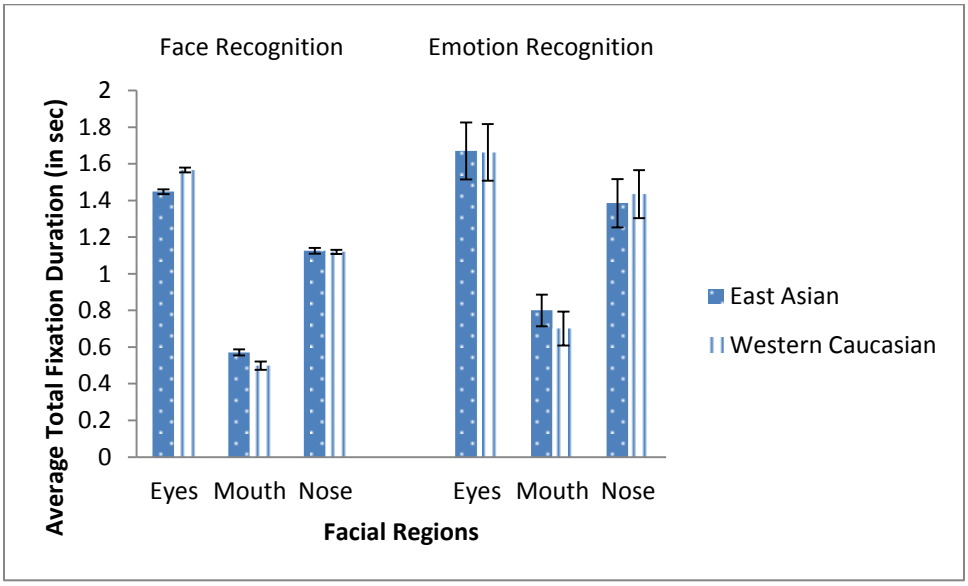


Figure C.1. Average total fixation duration for the eyes, mouth, and nose during the face recognition and emotion recognition tasks, in seconds. Error bars report standard errors of

mean. The fixation duration landing on the eyes was longer than the nose and mouth, and the nose longer than the mouth.

To examine the interaction between task and feature, separate Linear Mixed Models were conducted for each of the two tasks. Participant ID, face ID, and feature were included as factors (dependent variable = total fixation duration; fixed factor = feature of face). There was a main effect of feature for both face recognition and emotion recognition tasks, $F(2, 2397)=329.22$, $p<.001$ for the face recognition task and $F(2, 4197)=374.72$, $p<.001$ for the emotion recognition task. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for both tasks, the total fixation duration on the eyes was significantly longer than the nose ($p<.001$ for both tasks) and mouth ($p<.001$ for both tasks), and the nose longer than the mouth ($p<.001$ for both tasks). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and task were included as factors (dependent variable = total fixation duration; fixed factor = task). There was a main effect of task for the eyes, $F(1, 2198)=11.83$, $p=.001$, mouth, $F(1, 2198)=44.13$, $p<.001$, and nose, $F(1, 2198)=53.70$, $p<.001$. The total fixation duration on the eyes, mouth, and nose was significantly longer for the emotion recognition than face recognition task ($p=.001$ for the eyes and $p<.001$ for the mouth and nose).

To examine the trend towards an interaction between race and feature, separate Linear Mixed Models were conducted for each of the two races. Participant ID, face ID, and feature were included as factors (dependent variable = total fixation duration; fixed factor = feature of face). There was a main effect of feature for both races, $F(2, 3298)=292.82$, $p<.001$ for East Asian faces and $F(2, 3296)=373.73$, $p<.001$ for Western Caucasian faces. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for

both East Asian and Western Caucasian faces, the total fixation duration on the eyes was significantly longer than the nose (p<.001) and mouth (p<.001), and the nose longer than the mouth (p<.001). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and race were included as factors (dependent variable = total fixation duration; fixed factor = race of face). There was a main effect of race for the mouth, $F(1, 2198)=5.03$, $p=.025$, but not for the eyes, $F(1, 2198)=.92$, $p=.338$, and nose, $F(1, 2198)=.46$, $p=.5$. For the mouth, the total fixation duration on East Asian faces was significantly longer than Western Caucasian faces ($p=.025$).

Appendix D

Results for Total Fixation Duration in the Dynamic Face Recognition Task

For total fixation duration, there was a main effect of feature, $F(2, 4540.057)=394.634$, $p<.001$, and a trend towards a main effect of race, $F(2, 4540.057)=2.414$, $p=.09$ (see Figure D.1). For feature, post-hoc tests using Bonferroni Pairwise Comparisons showed that the total fixation duration landing on the nose was significantly longer than the mouth ($p<.001$) and eyes ($p<.001$), and the mouth longer than the eyes ($p<.001$). As for race, the total fixation duration for African faces was the longest, followed by East Asian then Western Caucasian faces; however all Pairwise Comparisons (Bonferroni corrected) failed to reach significance ($p=1$ for African and East Asian faces; $p=.122$ for African and Western Caucasian faces; and $p=.259$ for East Asian and Western Caucasian faces).

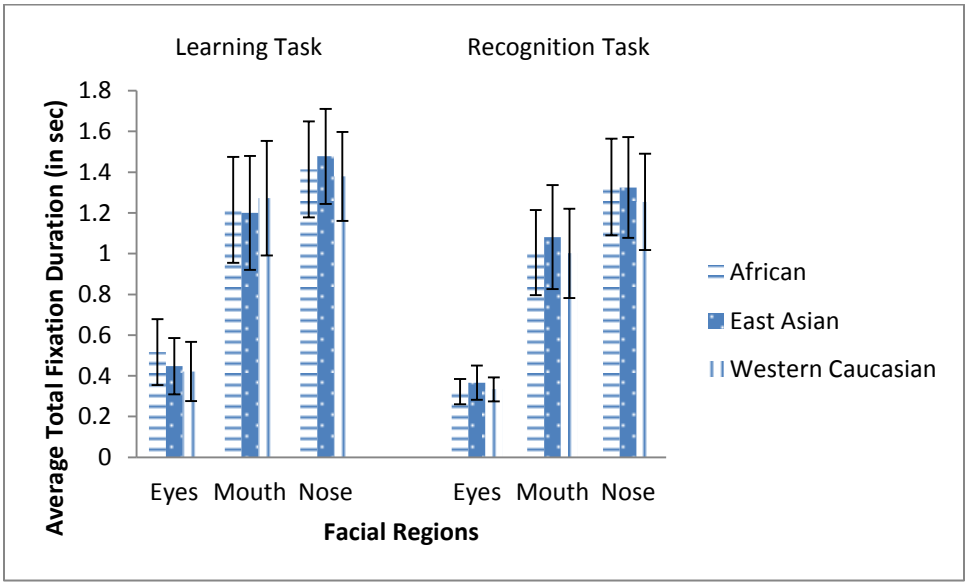


Figure D.1. Average total fixation duration for the eyes, mouth, and nose during the learning and recognition tasks, in seconds. Error bars report standard errors of mean. The total fixation duration for the nose was significantly longer than the mouth and eyes, and the mouth significantly longer than the eyes.

Appendix E

Results for Total Fixation Duration and Mean Fixation Duration in the Dynamic Emotion Recognition Task

For total fixation duration, there was a main effect of feature only, $F(2, 2457)=134.76$, $p<.001$, see Figure E.1. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that the total fixation duration for the mouth ($p<.001$) and nose ($p<.001$) was significantly longer than the eyes; however the total fixation duration for the mouth and nose did not differ significantly ($p=1$).

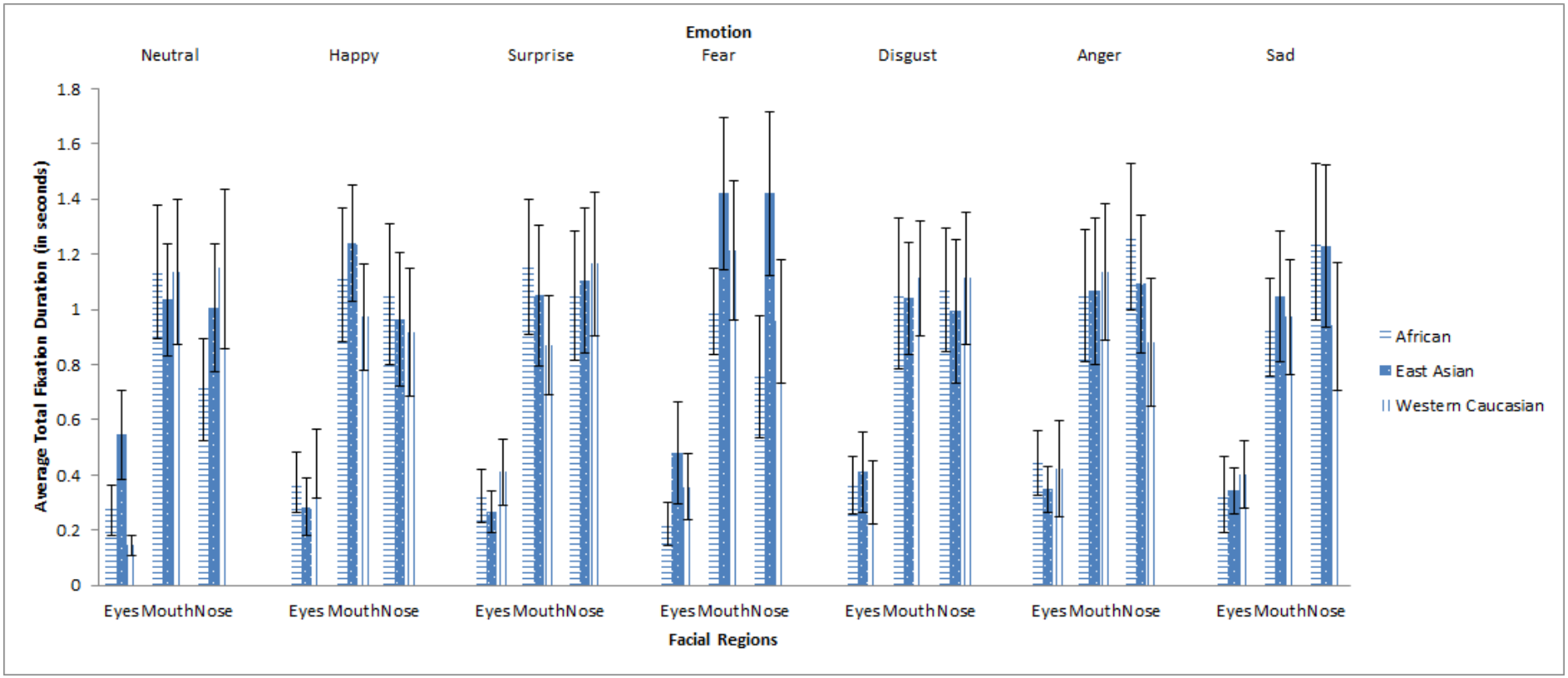


Figure E.1. Average total fixation duration for the eyes, mouth, and nose during the emotion recognition task, in seconds. Error bars report standard errors of mean. Participants fixated on the mouth and nose longer than the eyes.

For mean fixation duration, there was a main effect of feature only, $F(2, 1719.28)=15.27$, $p<.001$, see Figure E.2. For feature, post-hoc tests using Bonferroni Pairwise Comparisons revealed that the mean fixation duration for the mouth ($p<.001$) and nose ($p<.001$) was significantly longer than the eyes; however the mean fixation duration for the mouth and nose did not differ significantly ($p=1$).

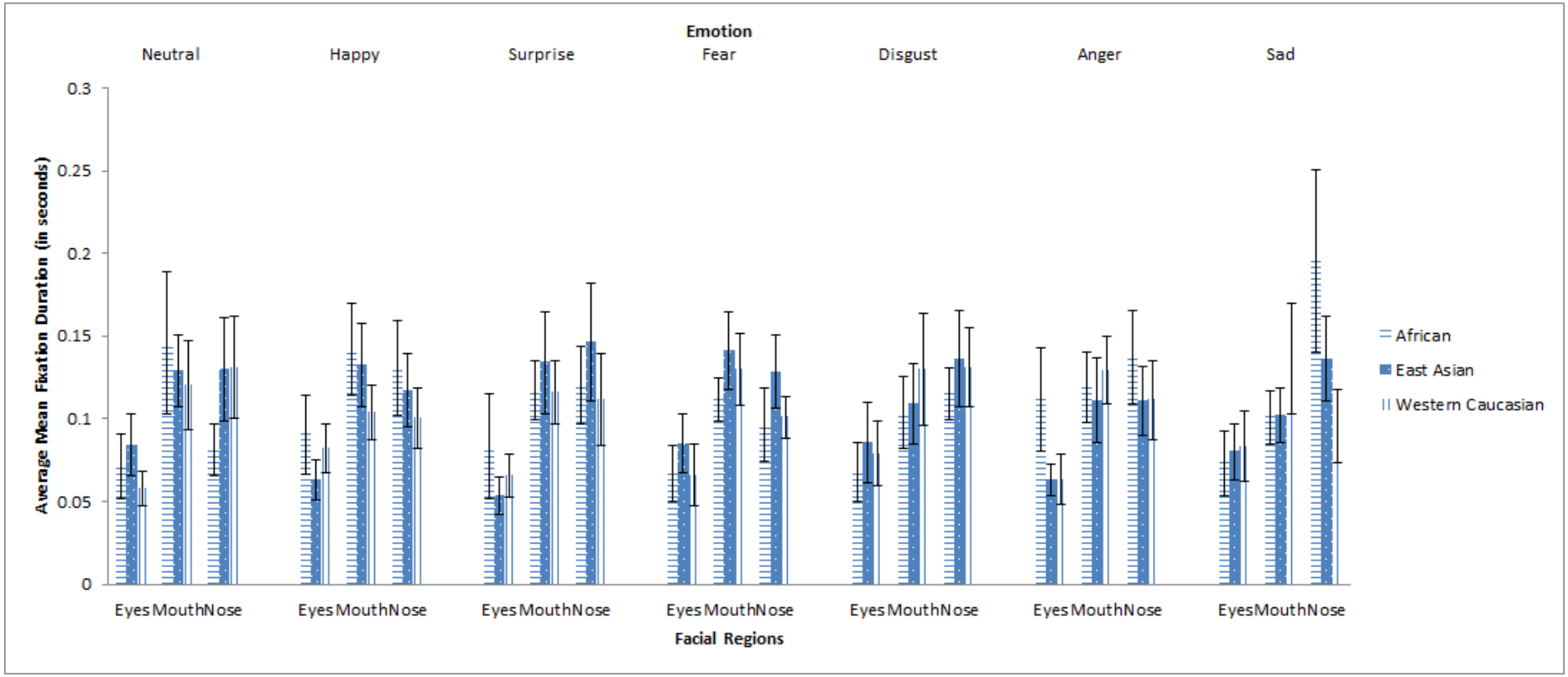


Figure E.2. Average mean fixation duration for the eyes, mouth, and nose during the emotion recognition task, in seconds. Error bars report standard errors of mean. The mean fixation duration for the mouth and nose was significantly longer than the eyes.

Appendix F

Results for Total Fixation Duration in the Dynamic Face Recognition and Emotion Recognition Tasks

For total fixation duration, there was a main effect of feature, $F(2, 613)=59.87$, $p<.001$, and an interaction between task and feature, $F(2, 613)=4.93$, $p=.008$, see Figure F.1. For feature, post-hoc tests using Bonferroni Pairwise Comparisons showed that the total fixation duration for the nose ($p<.001$) and mouth ($p<.001$) was significantly longer than the eyes; however the total fixation duration for the nose and mouth did not differ significantly ($p=.955$).

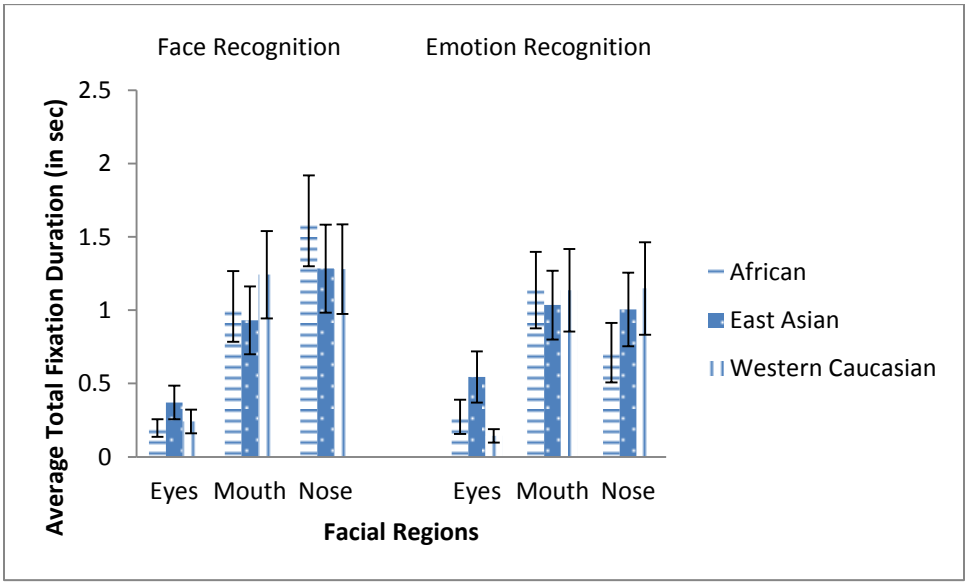


Figure F.1. Average total fixation duration for the eyes, mouth, and nose during the face recognition and emotion recognition tasks, in seconds. Error bars report standard errors of mean. For the face recognition task, the total fixation duration for the nose and mouth was longer than the eyes, and a trend towards longer total fixation duration for the nose than mouth, while for

the emotion recognition task, the total fixation duration for the mouth and nose was significantly longer than the eyes.

To examine the interaction between task and feature, separate Linear Mixed Models were conducted for each of the two tasks. Participant ID, face ID, and feature were included as factors (dependent variable = total fixation duration; fixed factor = feature of face). There was a main effect of feature for both dynamic face recognition, $F(2, 303)=35.07$, $p<.001$, and emotion recognition, $F(2, 357)=20.19$, $p<.001$, tasks. Post-hoc tests using Bonferroni Pairwise Comparisons revealed that for the dynamic face recognition task, the total fixation duration spent looking at the nose was marginally longer than the mouth ($p=.053$) and eyes ($p<.001$), and the mouth longer than the eyes ($p<.001$), whereas for the dynamic emotion recognition task, the total fixation duration spent looking at the mouth ($p<.001$) and nose ($p<.001$) were significantly longer than the eyes; however the total fixation duration on the mouth and nose did not differ significantly ($p=.774$). Further Linear Mixed Models were conducted for each of the three facial features. Participant ID, face ID, and task were included as factors (dependent variable = total fixation duration; fixed factor = task). There was a main effect of task for the nose, $F(1, 220)=7.35$, $p=.007$, but not for the eyes, $F(1, 220)=.54$, $p=.465$, and mouth, $F(1, 220)=.06$, $p=.811$. Post-hoc tests using Bonferroni Pairwise Comparisons showed that the total fixation duration landing on the nose was significantly longer for the dynamic face recognition than emotion recognition task.