Cultural similarities and differences in facial identity and expression processing

Xiaoqian Yan

PhD

University of York

Psychology

July 2016

Abstract

A range of research has shown an advantage for the perception of faces from same-race compared to other-races – the other race effect (ORE). Nevertheless, previous studies have focused on the ability to recognize facial identity. This thesis focuses on how information from the face that conveys expression is processed in own-race and other-race faces.

First, the cross-cultural processing of facial expressions of basic emotions in Western Caucasian and Chinese observers was investigated. Perceptual judgments of facial expressions were quite similar between cultural groups, but facial expressions from own-race faces were categorized more accurately than expressions from other-race faces. This is explained by differences in the recognition of lower region of the face. Facial expressions were processed in a holistic way, but there was no difference in the engagement of holistic processing of own-race compared to other-race faces.

Reliable own-group advantages were found in facial identity and expression processing with a free card-sorting task. However, there was also a large amount of cross-cultural consistency in response patterns.

Two core face-selective regions, the FFA and the OFA, were sensitive to changes in facial identities and expressions, but there was no difference in the magnitude of response to own- and other-race faces.

To summarize, this PhD thesis explored the cross-cultural processing of facial expression. Evidence showed significant differences in the perception of own-race and other-race faces, but these effects were generally small. The widely agreed opinion that the other-race effect is large is overstated.

Table of Contents

List of Figures

List of Tables

Acknowledgements

There are many people I would like to thank to during the past three years of my PhD life. It was difficult at the start to live and study in a foreign country, but I adapted to the new life very quickly with the help and encouragement of the people in the department and also my friends.

I would like to thank my supervisors Andy and Tim, from whom I learned how to be a real researcher. Thanks very much for your patience and invaluable advices to my studies. Their attitudes to science, the way of doing research have influenced me a lot and also made me who I am today.

I would also appreciate the support from all my lab colleagues, whose doors were always open to me. They have provided many help and suggestions throughout the study of my PhD.

I would also like to thank my parents, my fiancé, and also my friends, who have always been with me and believing in me during these years. I will always appreciate their unconditional love and supports. They are the most important spiritual supports in my life.

Author's declaration

Except where stated, all the work contained within this thesis represents the original contribution of the author under the supervision of Prof. Andy Young and Prof. Timothy J. Andrews. The work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Some of the materials presented within this thesis have previously been published in the following papers:

1. Yan, X. Q., Andrews, T. J., & Young, A. W. (2016a). Cultural similarities and differences in perceiving and recognizing facial expressions of basic emotions. *Journal of Experimental Psychology: Human Perception and Performance, 42*(3), 423–440.
2. Yan, X. Q., Andrews, T. J., Jenkins, R., & Young, A. W. (2016b). Cross-cultural differences and similarities underlying other-race effects for facial identity and expression. *The Quarterly Journal of Experimental Psychology, 69*, 1247–1254.

In addition, the following paper has been accepted for publication:

Yan, X. Q., Andrews, T. J., & Young, A. W. (accepted). Differences in holistic processing do not explain cultural differences in the recognition of facial expression. *The Quarterly Journal of Experimental Psychology*.

1. Introduction

Social signals from faces can be broadly divided into two relatively independent sources; static invariant characteristics and transient variant characteristics. From the invariant characteristics, people can tell the age, sex and identity of the faces; while from the transient changes of facial muscles especially from the eyes and mouth regions, people can tell a person's emotions, direction of attention and mouth movement. This introduction will explore how well people can process both the static and variant signals from the face (especially the facial identity and expression) and whether there are cross-cultural differences in perception. With increasing cross-cultural interaction and cooperation opportunities, people must communicate and function effectively with people from different cultural groups. Therefore, this is a topic of both theoretical and practical importance.

Studies have found that cultural differences exist in the way people process faces of their own-group versus other-group members. People are better at recognizing or remembering own-group faces, also known as the other-race effect (Meissner & Brigham, 2001). Similarly, studies also show that people are better at recognizing facial expressions of their own-group members (Elfenbein & Ambady, 2002b; Jack, Calder, & Schyns, 2012). However, no studies have systematically investigated the relationships of the cross-cultural processing between facial identities and expressions. It remains unclear whether this reflects a common bias in face recognition or independent processes, even though most accounts of face perception have postulated independent processing of facial identity and expression (Haxby, Hoffman & Gobbini, 2000; Calder &Young, 2005), since Bruce and Young (1986).

This thesis aimed to systematically explore cultural differences in the way people process facial identities and expressions of both their own-group and other-group members. First, this thesis investigated potential cross-cultural similarities and differences in perceiving and recognizing facial expressions of basic emotions, in order to address the recent debate concerning cross-cultural processing of facial expressions. Secondly, this thesis further explored potential factors (e.g., expression confusability, engagement in holistic processing) that might explain the cross-cultural differences in facial expression recognition. To measure the extent of the cross-cultural agreement and differences, the other-race effects for facial identity and expression were investigated at the same time with structure-matched tasks in the third study. And in the subsequent fMRI study, this thesis examined the neutral representation of the other-race effect in facial identity and expression processing in the human brain.

1.1 Cultural similarities and differences in processing facial expressions of basic emotions

1.1.1 Universality of facial expressions of basic emotions

One of the key issues of current debate in Psychology focuses on universality versus cultural relativity of emotional expression. During the last 40 years, the dominant position has been based on Ekman's (1980) interpretation of Darwin's (1872) proposal that a small number of basic emotions serve evolved biological functions, and that facial expressions of these basic emotions are universal across cultures. Consistent with the universal hypothesis, many studies have found that people can identify facial expressions of basic emotions portrayed by members of different cultures at above chance levels, even though there might be some variability across cultures (Biehl, et al., 1997; Ekman, 1972; Izard, 1971). Facial expressions posed by people in preliterate cultures were also found to be similar to expressions used by people from Western cultures (Ekman, 1972). McAndrew (1986) investigated American and Malaysian Chinese participants' recognition threshold of six facial expressions (anger, disgust, fear, happiness, sadness, and surprise). Their results found that even though there were minor group differences in the recognition of the anger expressions, cultural background had relatively little effect on their overall recognition accuracies and perceptual thresholds for the other five expressions.

However, the correct interpretation of such findings is still in debate. For example, the extent to which cross-cultural agreement is overestimated due to procedural factors such as the use of forced-choice recognition tests has been controversial (Ekman, 1994; Russell, 1994). In fact, even with six-alternative forced-choice recognition, the recognition rates ranged from 86% for Americans (Ekman, 1972) down to 53% for people in New Guinea (Ekman, Sorensen, & Friesen, 1969) when viewing American emotional expression images.

Ekman (1972) proposed a neurocultural theory of emotion, in which he argued that there was a one-to-one link between the emotion a person experiences and the facial expression the person poses. According to this model, people from all cultures express emotion in the same way in nonsocial settings. However, in different social settings people use certain display rules to control and manage the expression of emotions. These display rules can be different across cultures, and they are social norms that serve to intensify, diminish, neutralize, or mask emotional displays. But the forms of facial expressions of basic emotions themselves do not change substantially. Later on, Matsumoto (1989) extended this model by arguing that people across-cultures perceive emotions in the same manner, and the social norms actually work as a guidance to decode or explain the perceived emotions; people in different cultures may have different understandings of the same perceived emotion.

1.1.2 Culturally learnt facial expressions of basic emotions

Theories of cultural relativism have held sway through much of the twentieth century. From this perspective, it is thought that facial expressions are primarily culturally constructed and learnt (Elfenbein & Ambady, 2002b). Like the linguistic dialect, facial expressions of emotions may also have dialects that differ in ways of expression and explanation to result in potential cultural differences (Elfenbein & Ambady, 2002b). An own-group advantage has been found in many studies; recognition accuracy is higher when emotions are both expressed and perceived by members of the same cultural group. The own-group advantage is even found to be lower for groups with a closer geographical distance or having more cultural contacts with each other (Elfenbein & Ambady, 2002a, 2002b, 2003a, 2003b, 2003c).

Studies aiming at investigating cross-cultural differences in expression recognition or emotion intensity perception with fully balanced design (with different groups of participants identifying or rating faces of both their own-race and other-races) started in the early 90s (Matsumoto, 1992; Matsumoto & Ekman, 1989). Even though these studies have either found significant interactions involving observer culture and poser culture, they did not clearly decompose the interactions to clarify the relationship between the performance of observers and the stimuli ethnicity. In other words, whether or not people were better at recognition of facial expressions posed by their own race members was not fully reported. For example, Matsumoto (1992) asked American and Japanese participants to recognize and rate the intensity of facial expressions posed by members from their own- and other-cultural background. To decompose the significant interaction between observer ethnicity and stimuli ethnicity and emotion, they only noted that Americans were better than Japanese at identifying certain expressions, like anger, disgust, fear, and sadness, but did not mention whether or not Americans were better than Japanese at identifying expressions posed by their own-race members.

Recent studies have found that cross-cultural differences in facial expression recognition may be driven by people from different cultural backgrounds tending to focus on different facial signals. By asking participants to judge the intensity of emoticons or facial images that were generated by combining eyes and mouth regions from happy and sad faces, Yuki and colleagues (2007) found that East Asian participants give more weight to the eyes of the emoticons and facial images compared to Western participants, whereas Western participants give more weight to the mouth region of the faces when rating expressions in comparison with East Asian participants. Jack and colleagues (2009) used the eye-tracking technique to determine the key features needed to recognize facial expressions by Western Caucasian and East Asian participants. They found that when recognizing facial expressions Western participants fixated features in the eye and mouth regions, whereas East Asian participants mainly fixated the eye region, which was consistent with Yuki and colleagues' (2007) finding.

Even though eye fixation patterns can indicate increased interest in certain regions of the face, they do not rule out the possibility that features from less fixated regions are none the less attended or encoded. In a further study, Jack et al. (2012) used the reverse correlation method to determine whether the differences in fixation patterns reflect different mental representations of key expressive features. They asked participants to make six-alternative forced-choice identification of the facial expressions of 12,000 highly ambiguous stimuli that were derived by adding pixel-based white noise to a neutral face base image. They then averaged the white noise templates associated with each categorical judgement by a participant, to capture that participant's internal representation of facial features of each facial expression. In line with the eye movement findings (Jack et al., 2009), Jack et al.'s (2012) results showed that Western participants used information from the eye and mouth regions to represent facial expressions internally, whereas East Asian participants relied largely on the eye region, including the eyebrows and eye gaze direction.

Jack et al. (2009, 2012) therefore provided evidence supporting differences between Western and East Asian participants' representations of facial expressions, and specifically highlighted differences in the use of the eye and mouth regions. However, there still are some aspects of the techniques they used that need further confirmation. In particular, even though reverse correlation can in theory potentially capture any of the communicative signals participants might seek, in practice the potential variety of facial expression cues means that stimuli created by adding pixelated noise to a neutral face are unlikely to contain sufficient examples of all possible facial signals (Freeman & Ziemba, 2011). Therefore, Jack et al.'s (2012) findings concerning how culture can finely shape the internal representation of facial expressions might be influenced by these constraints of the reverse correlation method they used.

One study also finds that there are cross-cultural differences in recognition of facial expressions at subdued intensity levels: subtle, low, and moderate (Zhang, Parmley, Wan, & Cavanagh, 2014). There was an own-group advantage at the low and moderate intensity level for anger expressions, and the moderate intensity level for sad expressions. But at the low expression intensity, American observers were more sensitive to both the Western

Caucasian and Chinese expressions than Chinese observers. Their results indicate that cultural differences even exist in the recognition of subdued facial expressions, and the recognition of facial expressions may in part influenced by cultural and social norms. Compared with Western Caucasian individuals, Eastern Asians learn to control and hide their negative emotions to avoid social and interpersonal harmony (Matsumoto, 1989; Matsumoto & Ekman, 1989), which could further make them less sensitive to the negative low intensity expressions.

As well as investigating cultural differences in recognition of facial expressions, researchers also investigated the cultural differences in judgments of emotion valance and intensity. Studies have found that people tend to give higher ratings to expressions posed by their own-race members versus other-race members (Sneddon, McKeown, McRorie, & Vukicevic, 2011; Zhu, Ho, & Bonanno, 2013). However, these studies yield mixed results. Soto and Levenson (2009) asked their African American, Chinese American, European American and Mexican American participants to rate the valance and intensity of the emotions when they were watching several videos, but they did not find the own-group advantage. However, the study did not have a fully balanced design because all the participants were American. Their perception of emotions might have been affected by the consistent social display rules.

1.1.3 Current debates concerning the differences between static and dynamic expression stimuli

Recent studies have been questioning the validity of using static expression stimuli to investigate human processing of facial expressions, even though static images have been widely used in many studies. Researchers have argued that using static images not only lacks ecological validity, but also limits our understanding of the meanings of facial activities. Compared with static images, dynamic expressions can convey information not only about the presence of an emotional state, but also about the unfolding and ending, temporal sequence, speed, and also the quality information (Krumhuber, Kappas, & Manstead, 2013).

However, studies have found inconsistent results regarding the advantages of dynamic expressions over static expressions. Wehrle, Kaiser, Schmidt, and Scherer (2000) found that dynamic synthetic expressions with entirely posed facial muscle configurations increased the overall recognition accuracy and reduced confusions between different emotion categories compared to static images. Recent neuroimaging studies also find that more broad regions of the parahippocampal gyrus are activated by dynamic versus static images, especially in the right hemisphere, including the amygdala, inferior occipital gyri, middle temporal gyri, fusiform gyri, inferior frontal gyri, and the orbitofrontal cortex (Arsalidou, Morris, & Tayloer, 2011; Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004; Trautmann, Fehr, & Herrmann, 2009). These regions are closely related with social- and emotion-related information processing. By measuring the magnitude of response in several core face-selective regions according to Haxby and colleagues' neural model of face perception (2000) (find more details in section 1.4.2) to both dynamic and static stimuli, Pitcher and colleagues (2011) found that responses to dynamic movies of faces were not significantly different from the responses to static images in the right fusiform face area (FFA) and the right occipital face area (OFA), but the responses in the right posterior superior temporal sulcus (STS) were nearly three times as strong to dynamic faces as to static faces. A further TBS (thetaburst transcranial magnetic stimulation) study by Pitcher and colleagues (2014) showed that dynamic and static facial aspects were processed via dissociable cortical pathways in human brain.

Nevertheless, when comparing participants' recognition accuracy and response time across both static and dynamic facial expressions, Fiorentini and Viviani (2011) did not find a dynamic expression advantage. Johnston and colleagues (2013) found that compared to a sex discrimination task, emotion discrimination judgements involved more widespread regions in the occipito-temporal, parietal and frontal cortex and these regions were activated equally by both dynamic and static facial expressions, but regions of the inferior frontal gyri and supplementary/pre-supplementary motor areas were more activated to static faces. Krumhuber et al. (2013) further summarized in their review paper that dynamic advantages are only evident when information conveyed by static images is insufficient or unavailable, or when people have neurological or developmental disorders (e.g., brain damage or autism), or when the intensity of the expressions is subdued.

Even though there have been differences between dynamic and static facial expressions, this does not exclude the validity and benefits of using static expression images as experimental stimuli. First, static expression images can catch the apex of expressed emotions, which is sufficient for the recognition of many facial expressions. Secondly, based on the fact that a lot of previous studies have used static images to investigate expression processing, it is appropriate to use static images as well in some studies in order to make a comparison or to make further extensions based on previous research findings. Using static images can thus avoid bringing in more confounds to the results. In the cross-cultural face perception field, most of the studies still tend to use static faces as stimuli. However, in Jack and colleagues' (2012) study, three-dimensional facial movement stimuli were randomly created and participants were required to make forced-choice expression judgments on these random facial animations. Their results still showed that Eastern Asians and Western Caucasians represented dynamic facial expressions very differently, consistent with their findings with static face images.

1.2 Cultural differences in processing facial identities

Apart from the own-group advantage in facial expressing recognition, there is also an own-group advantage, usually called the other-race effect (ORE), in facial identity recognition tasks. In a meta-analysis across 39 studies, Meissner and Brigham (2001) confirmed the strong and reliable other-race effect in face recognition; people were more accurate at remembering individuals from their own cultural groups than other cultural groups. In addition, this ORE has also been replicated in applied eyewitness identification studies

(Brigham, Bennett, Meissner, & Mitchell, 2007; Chance & Goldstein, 1996). Eyewitnesses are more likely to misidentify suspects of another cultural group compared to those of their own groups. However, most of the studies reporting the ORE have been based on the fact that participants were better at remembering faces of their own-group members.

Recent studies have also showed that the ORE exists at the perceptual level. Lindsay, Jack and Christian (1991) first suggested that the ORE might actually reflect the perceptual skills specific to identify faces of own-group members. In their study, groups of White and African American participants were required to do a match-to-sample task, in each trial of which a White or African American face was shown to the participants at first followed by a test pair comprising both the target face and a matched foil. The results showed that White participants performed more poorly on trials involving matching African American faces than on trials involving White faces, even though no such differences were found among African American participants. In a similar same/different sequential matching task, Walker and Tanaka (2003) generated continua of images through morphing between an East Asian and a Western Caucasian face (see Figure 1.1). In each trial, participants were presented with an East Asian or Western Caucasian parent face followed by either the same parent face (a ‘same’ trial) or an Asian-Caucasian morph face (a ‘different’ trial). The morph faces were always morphed with 90%, 80%, 70%, 60%, or 50% contributed from the same parent faces present at first and the remaining percent contribution from the opposite-race parent face. Their task was to judge whether the two faces they saw sequentially were physically the same or different. Results showed that East Asian participants were better able to discriminate the East Asian parent faces among the morph faces, whereas Western Caucasian participants were better able to discriminate the Western Caucasian parent faces among the corresponding morph faces. Following up the work of Bruce et al. (1999) which asked participants to find a target face in a line-up of 10 face images, Megreya and colleagues (2011) found that both British and Egyptian participants were worse at matching other-race faces than own-race faces, confirming that difficulty in perceptual encoding of unfamiliar faces contributes substantially to the other-race effect.

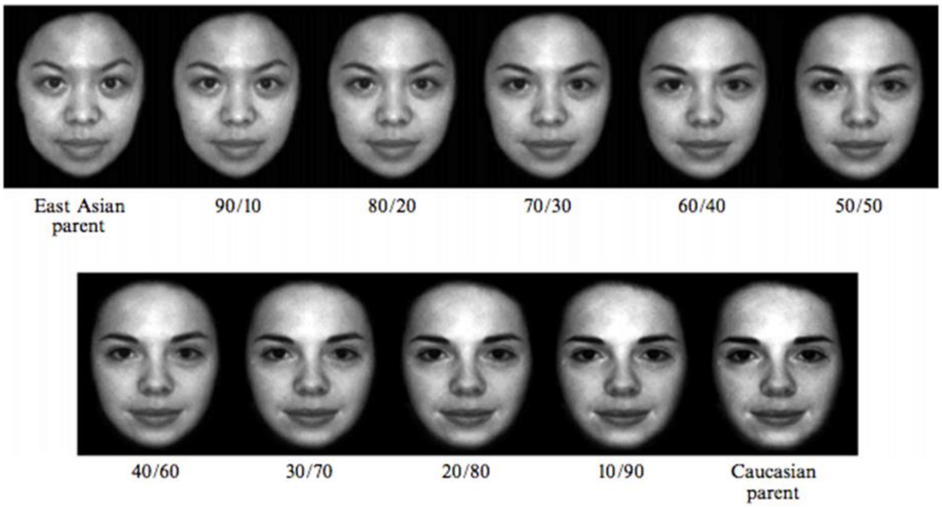


Figure 1.1 Example of the continuum of morph faces in Walker and Tanaka's (2003) study

1.3 Underlying mechanisms in the cross-cultural differences

1.3.1 Current theories explaining the own-group advantage in facial expression processing

Studies have suggested that the own-group advantage in recognizing facial expressions can be caused by cultural differences in display and decoding rules regarding facial expressions

(Ekman, 1972; Matsumoto, 1989). Some cultures, like the Japanese culture, impose social norms to avoid cases when expressing or understanding certain emotions may be disruptive to social harmony. People thus might be better at understanding emotions expressed in a way matching their own. Even though several nonverbal behaviors are likely to descend from a common "nonverbal 'language'", some of the specifics of this language still vary across different cultural groups (Rosenthal et al., 1979). Cultural variations in facial expression identification can also be attributed to variations in the way of encoding across cultures (Elfenbein & Ambady, 2002b; 2003c). Recent studies by Jack and colleagues (2009, 2012) also support the encoding difference theory, suggesting that the differences between cultural groups are driven by people from different cultural backgrounds paying attention to different facial signals when processing facial expressions. For example, East Asian participants mainly focus on the eye region when processing facial expressions, whereas Western Caucasian participants fixate more evenly across the eye and mouth regions. From this perspective, the cross-cultural differences reflect underlying differences in mental representations resulting from differences in the attended regions of the face.

The cross-cultural differences in expression recognition can also be modulated by differences in motivation. People may be less motivated to understand emotions expressed by members of other cultural groups (Kilbride & Yarczower, 1983; Markham & Wang, 1996). By arbitrarily allocating participants into different groups, Young and Hugenberg (2010) found an own-group identification advantage even with the race of face stimuli and participants held constant. However, the difference in motivation cannot be the only factor to explain the own-group advantage. In a study using face stimuli among multiple Western Caucasian groups, Elfenbein and Ambady (2002b) still found an own-group advantage, even though it was very hard for the participants to determine the cultural background from each expression image because the differences among them were minimal.

1.3.2 Current theories explaining the other-race effect in facial identity processing

Theories to explain the other-race effect for identify recognition mainly include the perceptual learning theories and the social cognitive theories. The shared main idea of the perceptual learning theories is that people have different experiences encoding own-race and other-race faces. Different levels of experience then lead to different recognition of own-group comparative to other-group faces (Chiroro & Valentine, 1995; Hancock & Rhodes, 2008; Meissner & Brigham, 2001; Walker & Tanaka, 2003; Walker & Hewstone, 2006). With increased contact with own-race members, people's visual systems become more efficient at discriminating among faces of their own race versus those of other races (Furl, Phillips, & O'Toole, 2002). Sangrigoli and colleagues (2005) found that after being adopted by Western Caucasian families since they were young (3 to 9 years old), their Korean origin participants actually performed better at recognizing Western Caucasian faces than East Asian faces, indicating that face recognition ability can be modulated by experiences with faces. Walker et al. (2007) also found that other-race experiences played an integral role in the way people process own-race versus other-race faces from even the early perceptual stages of processing. However, not all studies have found the effect of social contact on the recognition or perception of faces, and the meta-analytic study by

Meissner and Brigham (2001) found that social contact with other race faces could only explain 2% of the variance in the ORE.

The social cognitive theories emphasize the effect of social categorization of own-race and other-race faces. For example, people are less motivated to process faces they see as belonging to another group (Hugenberg, Young, Bernstein, & Sacco, 2010; Bernstein, Young, & Hugenberg, 2007; Sporer, 2001). Levin (2000) argued that differences in categorization processes for own-race versus other-race members can lead to a quick categorization of other-group faces at the expense of encoding individual information. People tend to categorize members of their own race at a subordinate level of the individual (e.g., Bob, Joe) and categorize other race members at the basic level of race (e.g., Western Caucasian, East Asian). By training Western Caucasian participants to categorize African or Hispanic faces at either the subordinate individual level or the basic level, Tanaka and Pierce (2009) found that training at the subordinate individual level led to better post-training face recognition tests and also a larger amplitude in the N250 component, which is an indicator of familiar face recognition (Schweinberger, Huddy, & Burton, 2004; Schweinberger, Pickering, Burton, & Kaufmann, 2002), than training at the basic level. However, by presenting morphed ambiguous-race faces in either Western Caucasian or East Asian face contexts to form race categorization, Rhodes and colleagues (2010) did not find the corresponding own-group advantages in either the recognition or perception of the identical morphed faces.

Hugenberg and colleagues (2010) proposed an integrative Categorization-Individuation Model, in which they argued that social categorization, initial motivation and perceptual experiences worked together to create the other-race effect. This model holds that there are two different ways of face processing: categorization and individuation. Categorization means classifying exemplars into different groups along shared dimensions, which occurs quickly, automatically and simultaneously when encountering faces in most situations; while after categorization, individuation occurs to discriminate among exemplars within a category. The individuation process requires attention to facial features for identity processing. According to this model, categorization is first activated upon the presentation of a target face, and this can cause other-race faces to be processed more homogeneously. Perceiver motivation is another factor to direct attention to categorization and individuation processing. Personal motivation can even direct attention to identity processing of other-race faces, even though in most cases, identity of other-race faces seems irrelevant. Individual experience is another factor that affects the other-race effect, but its effect on facial identity processing works together with different motivation to individuate.

1.3.3 Differences in the engagement of holistic processing as another factor to explain the cross-cultural differences?

Studies have claimed that humans have specific ability to process faces versus other objects. The specific ability has been attributed to the configural processing of faces because Yin (1969) found that the ability to recognize faces reduced substantially when faces were upside down compared to recognition of other objects. The term configural processing here indicates the processing that involves perceiving spatial relations among features of a stimulus, like a face. However, studies have shown that configural processing is not specific to faces, as after training or after developing as an expertise, people can also apply configural processing to other categories of objects, especially those as visually homogeneous as faces (Bruyer & Crispeels, 1992; Diamond & Carey, 1986; Gauthier, Williams, Tarr, & Tanaka, 1998; Gauthier, & Tarr, 2002).

Maurer, Le Grand, and Mondloch (2002) suggested that the idea of configural processing at least comprises three different types: (1) processing of first-order relations, that is to see a stimulus as a face because it has two eyes above a nose, which is above a mouth; (2) holistic processing, that is to perceive a face as a whole or a gestalt; (3) processing of second-order relations, that is to be sensitive to certain distances among features. These three types of configural processing can be distinguished by different behavioural tasks (Maurer et al., 2002).

According to Maurer et al. (2002), sensitivity to the first-order configural relations allows people to process faces specifically versus other objects. Preferences for face-like stimuli appear even in newborns (Johnson et al., 1991; Mondloch et al., 1999). By presenting participants with upright or inverted stimuli of faces, houses, or other objects, Yin (1969) found that people’s ability to identify faces declined dramatically versus other objects, when they were presented upside-down. This particular face inversion effect has been replicated in many subsequent studies (Valentine, 1988), reflecting the important role of first-order relations in the perception of faces.

If we say the first-order relations with two eyes on top of a nose and the nose above a mouth make faces specific to other objects, studies also found that people process faces in a more holistic way than a feature-based processing (here I am not saying people only process faces holistically, people also need the feature information to identify faces). The holistic processing of faces has been widely established by two tasks, the composite-face task and the whole-part task. In the composite task, Young, Hellawell and Hay (1987) combined the upper- and lower-half faces of different famous identities either in a vertically aligned way or misaligned way (by positioning the nose from the lower-half faces close to one edge of the upper-half faces), before asking participants to name the identities from either half of the face. The results came out to be that the noncomposite faces were named faster than the composite faces for either half of the faces, indicating that the perception of a novel face in the composite conditions, attributed to the holistic processing of faces, interfered with the identification of the part of the faces. In the part-whole task, Tanaka and Farah (1993) asked participants to identify faces either from isolated features of two learned faces (e.g., eyes, nose, mouth), or from two full-set faces that differed only in one individual feature that were tested in the isolated condition. They found that the same face part was significantly better identified in the whole face condition compared with the isolated part condition, while the whole-face advantage disappeared in scrambled faces or in inverted faces, indicating that faces are mentally represented in a more holistic way.

Studies have also suggested the ability to use differences in the second-order relations among faces is critical for face perception (Diamond & Carey, 1986). To examine the second-order relations, set of face-like images have been created by varying only the spacing of different facial features (e.g., eyes, nose, and mouth) with the features identical in all faces. Studies have found that inverting the face stimuli could affect the ability to discriminate faces when they differed only in second-order relations much more than faces that were different only in features (Freire, Lee, & Symons, 2000; Le Grand et al., 2001). Although other studies also found inconsistent results suggesting that the second-order relations and facial features have similar effect on face identification (Yovel & Duchaine, 2006; Yovel & Kanwisher, 2004a, 2004b), these studies all indicate that the ability to discriminate differences in second-order relations or features among faces plays an important role in face recognition. However, one recent study challenged the importance of the second-order relations hypothesis in face recognition, especially in familiar face recognition. In a series of three experiments, Sandford and colleagues (2014) rescaled the length and width of face images to random values between 50% and 200% of their original, and then asked participants to adjust the image window until the image looked right. Their results did not find the familiar face advantage compared to unfamiliar faces, which was originally predicted by the authors because participants would not know the detailed spatial differences between features of unfamiliar faces. The results hence indicated that the subtle differences in spatial relations between facial features are not important, or are not used at all, in recognizing familiar faces.

Now, let us return back to the main question of this section. I will only focus on the holistic processing of faces in this thesis, and investigate whether the cross-cultural differences in face processing are partly driven by the differences in the engagement of holistic processing to different race of faces. Even though the perceptual learning theories and the social cognitive theories emphasize different factors causing the ORE in face processing, they actually all agree that the underlying mechanism of this effect is that own-race faces are processed in a relatively more holistic way than other-race faces. According to the perceptual experience theories, the representations underlying the holistic processing of faces are coarsely applied to faces with different races, but extensive perceptual experience will promote relatively more holistic processing of own-race faces (Michel, Caldara, & Rossion, 2006a; Michel, Rossion, Han, Chung, & Caldara, 2006b; Tanaka, Kiefer, & Bukach, 2004; Rhodes, Hayward, & Winkler, 2006). The difference in holistic recognition between own- and other-race faces can also mirror the difference in people's relative experience with own-group and other-group faces (Tanaka et al., 2004). According to the social cognitive theories, own-race faces are processed at the individual level and require more holistic processing, while other-race faces are processed only at the category level with less holistic processing. Michel, Corneille and Rossion (2007) found that identical morphed faces turned out to be processed more holistically when they were categorized as own-group faces versus other-group faces.

The above mentioned studies indicate that differences in the engagement of holistic processing play an important role in the ORE, but it cannot be the only contributing factor.

Studies also showed that sensitivity to component changes was greater to own-race faces compared with other-race faces (Rhodes et al., 2006; Hayward, Rhodes, & Schwaninger, 2008), indicating that the advantages at recognizing own-race faces are actually caused by a general facilitation in different forms of face processing.

Therefore, is it possible that the own-group advantage to recognize facial expressions is also caused by the difference in the engagement of holistic processing? As well as the fact that a number of studies have linked the ORE in the recognition of facial identity (rather than expression) to holistic processing (Tanaka et al., 2004; Michel et al., 2006a; Michel et al., 2006b), recent studies also indicated that there might be differences in holistic processing to facial expressions between different cultural groups. As mentioned above, Jack et al.'s (2012) finding that East Asian participants mainly use information from the eye region to recognize facial expressions predicts that holistic processing of expression should be reduced in comparison to Western Caucasian participants. However, for Western participants it is well-established that facial expressions are perceived holistically, with information from the mouth region modifying the interpretation of information from the eye region and vice versa. The most well-known demonstration of holistic processing involves a facial expression variant of the face composite paradigm devised by Young et al. (1987). Calder and colleagues (2000) created images that combined the upper half of one expression with the lower half of a different expression. They found that participants were slower at identifying expressions from either the upper or the lower part of these images when the two half parts were presented in a face-like aligned composite format than when the same parts were presented in a misaligned format that was not face-like. This effect has been replicated in other studies of Western participants (Calder & Jansen, 2005; Flack et al., 2015; Tanaka, Kaiser, Butler, & Le Grand, 2012; White, 2000). It is interpreted as indicating that holistic perception of the face-like aligned composite stimuli makes it difficult for participants to ignore information from the irrelevant part of the image (i.e. to ignore information from the lower half when classifying the upper half, or vice versa). In contrast, because the misaligned stimuli do not create a face-like configuration, they are not susceptible to this holistic interference.

However, no studies have yet examined the cross-cultural differences in the holistic processing of facial expressions. This gap in the literature might be driven by the fact that unlike the prevalent and widely established other-race effect in facial identity processing, there are still controversies concerning cultural differences in facial expression recognition. To address this, a systematic investigation of cross-cultural similarities and differences in the perception and categorization of facial expressions of basic emotions is needed, before investigating the possible cultural differences in the engagement of holistic processing of facial expressions.

1.4 Neural mechanisms underlying the cross-cultural differences in facial expression and identity processing

1.4.1 The functional model of face processing

To understand the neural mechanisms of how people from different cultures process own-race and other-race faces, we must first understand how we human beings process faces.

According to a widely discussed functional model for face processing, Bruce and Young (1986) suggested that seven different kinds of information can be distinguished from faces, which are labelled pictorial, structural, identity-specific semantic, visually derived semantic, name, expression, and facial speech codes (see Figure 1.2). In everyday face processing, incoming information is first structurally encoded based on the perceptual input and then involves more abstract identity-specific semantic processing of the face before the person's name is retrieved.

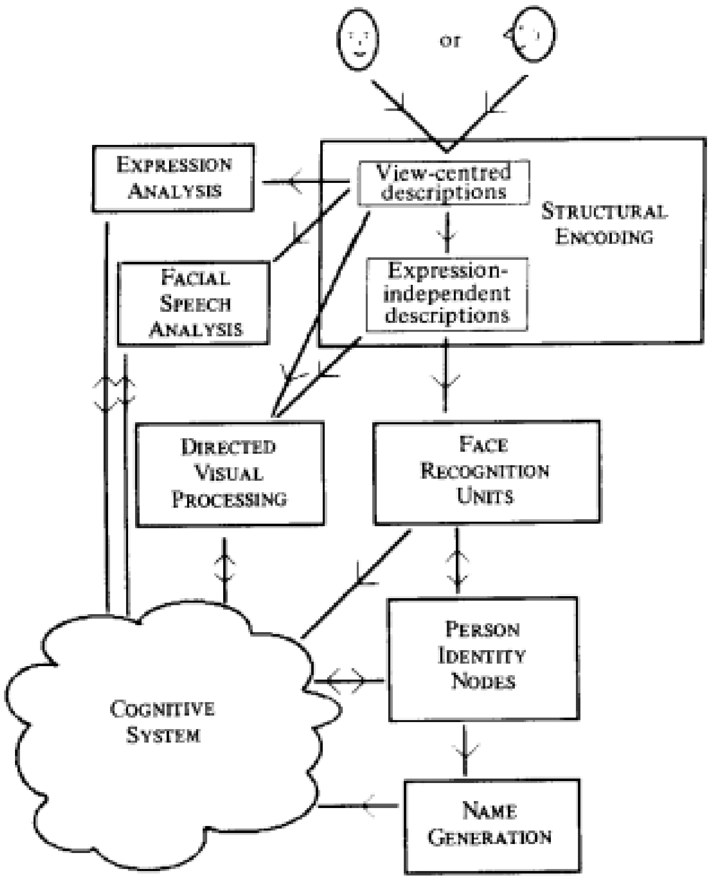


Figure 1.2 Functional model of face perception by Bruce and Young (1986)

To be more specific, the "structural encoding", represented as the first step of face processing in this model, comprises two separate representations: "view-centred" and "expression-independent" processing. The low level visual information is first extracted from the target face in the "view-centred" unit. The representation in this stage is influenced by the size, shape, texture, or orientation of the face. The "expression-independent" processing thus combines information flow from the "view-centred" stage and represents the abstract configuration of the target face as a whole. The representation of the face in this "expression-independent" stage is independent of the facial expression and appearance according to the model. After the first two stages of "structural encoding" of the target face, the newly formed representation of the face will compared with the already stored knowledge of this face held in the "face recognition units". If the comparison is evaluated as similar enough, the target face is recognized. And then the identity-specific semantic knowledge about the person is accessed via the "person identity nodes". After the effective retrieval of semantic information, the name of the person is then accessed in the "name generation" unit.

The information stored in the "face recognition units" contains only the visual appearance of the target face, which is stored in the long-term memory. While the "person identity nodes" act as a multimodal gateway accessible from including the person's face or voice. Distinctive semantic information is also associated with other relevant information through the "cognitive system". For example, the information in the "expression analysis", "facial speech", and "directed visual processing" can all flow into the "cognitive system" and link to the "person identity nodes".

A key claim of this functional model is that recognition of identity-related and identity-related information, such as facial expression, are relatively independent from each other. This point has been supported by many studies. For example, Young, McWeeny, Hay and Ellis (1986) found that recognition of facial expression was not influenced by the familiarity of the face. The independent processing of identity and expression was also found for the composite face effect (Calder et al., 2000). Neuropsychological case studies also found relatively independent recognition of facial expression and identity in patients with discrete brain lesions (Young, et al., 1993; Baudouin & Humphreys, 2006).

However, the idea of independent processing of facial expression and identity remains controversial. For example, Schweinberger and Soukup (1998) found that changes in irrelevant identity had an effect on recognition of expression, while irrelevant changes in expression did not affect recognition of identity. Later studies also found adaptation aftereffects for facial expression were larger than the aftereffects when the adapting and test stimuli had different identities (Campbell & Burke, 2009; Ellamil, Susskind, & Anderson, 2008; Fox & Barton, 2007; Fox, Oruç, & Barton, 2008). In a series of experiments with Garner's speeded-classification task, in which participants were required to make speeded classifications of either the expression (smiling or angry) or identity (Person A and Person B) of faces, Ganel and Goshen-Gottstein (2004) showed that Garner interference was found both from identity to expression, and vice versa. Garner interference between identity and expression was larger for familiar versus unfamiliar faces.

The inconsistent results mentioned above might be driven by differences in task and also in experimental parameters. In addition, to find a clear relationship between identity and expression, two tasks used must be matched on their level of difficulty, or else participants' better performance in one task might be mistaken as a separate processing instead of the relative ease of that task (Calder & Young, 2005).

The functional model by Bruce and Young (1986) described the first stages of recognition a face as involving the construction of a structural representation of the face. However, we can get some hints about the representation of own-race and other-race faces from this model. Low level visual information is first extracted from the target face in the "Structural encoding" stage, in which the orientation, size, and colour of the target face are encoded. As faces of different races can differ in average shape, skin colour and also hair colour (Bruce & Young, 2012; Farkas, Katic, & Forrest, 2005), these cues could be captured in the "Structural encoding" stage, as well. This model also allows the possibility that face representations will be constructed with different contributions of different types of information, as a function of expertise. People from different cultural backgrounds may attend to different areas of a face to maximally match the profile saved in the "Face recognition units", for example, studies have found that East Asians focus more on the central region of the face, whereas Western Caucasians fixate both the eye region and the mouth region during an old/new recognition task (Blais, Jack, Scheepers, Fiset, & Caldara, 2008). People's expertise can therefore lead to better recognition of own-race faces than other-race faces.

1.4.2 The neural mechanisms in face processing

Studies have shown that face-specific activations can be seen in many regions of the brain, including a region in the inferior occipital gyrus (occipital face area, OFA), the lateral side of the mid-fusiform gyrus (fusiform face area, FFA), the superior temporal sulcus (STS), and also the amygdala (see Figure 1.4 from Harris et al., 2014b). However, the most robust and dominant face-specific region is the FFA (Kanwisher, McDermott, & Chun, 1997). Since the 1990s, studies started to show that the fusiform region responds more strongly to faces than letter strings and textures (Puce, et al., 1996), flowers (McCarthy, Puce, Gore, & Allison, 1997), and other type of stimuli like objects, houses, and hands (Kanwisher, et al., 1997). Studies also showed that there was strong fMRI adaptation in the FFA to identically presented faces versus new faces (Yovel & Kanwisher, 2004a; Rotshtein, Henson, Treves, Driver, & Dolan, 2005). To rule out the possibility that the FFA is just highly sensitive to the configuration/spacing relations, Yovel and Kanwisher (2004a) examined whether or not participants process houses in the same way to faces by asking them to judge pairs of faces or house that differed in either the spacing distances among features or the shape of features. The FFA activation to faces turned out to be three times as strong as the FFA activation to houses, indicating that the FFA is mainly engaged in face processing. Studies also find that the FFA responds much stronger to faces that are consciously perceived, compared with the responses activated with non-conscious perception (Jiang & He, 2006; Williams et al., 2004).

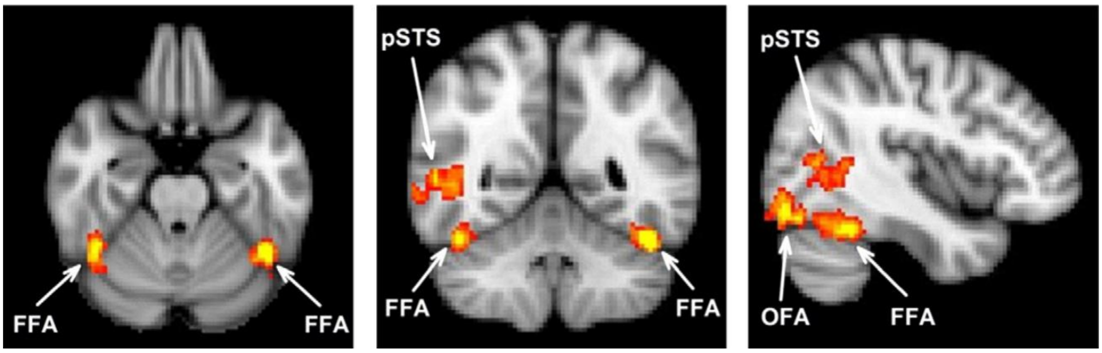


Figure 1.3 Average location of the FFA, the OFA, and the STS (Harris et al., 2014b)

Compared to the FFA, which is found to be more sensitive to changes across face identities rather than physical changes, studies have found that OFA is more sensitive to physical changes between faces, regardless of changes in identities (Fox, Moon, Iaria, & Barton, 2009; Rotshtein et al., 2005; Yovel, et al., 2005). For example, Rotshtein et al. (2005) found that the OFA responded similarly to two morphed faces that differed physically regardless of whether or not participants perceived the two faces are similar. On the contrary, the FFA only responded to different identities of faces rather than to physical differences. Yovel and Kanwisher (2005) found that the OFA responded similarly to both upright and inverted faces, and the different responses in the OFA to upright versus inverted faces did not correlate with the behavioral face inversion effect. Compared with the OFA, the FFA responded strongly to upright faces rather than inverted faces, and these differences in the FFA activation were correlated with the behavioral inversion effect. Upright faces were represented by face-specific mechanism whereas inverted faces were processed by both face-specific and object-specific mechanisms (Pitcher et al., 2011).

The Superior Temporal Sulcus (STS), along the full length of the temporal lobe, is one of the longest sulci in the brain. It is considered that the STS plays a central role in social perception and cognition. As well as being sensitive to the perception of biological movement, such as the movement of the whole human body, the hand, and the eyes and mouth (Puce, et al., 1998; Calvert, et al., 1997; Grèzes, et al., 1998; Bonda, et al., 1996), subregions of the STS are also activated during the perception of still images of face and body, indicating that it is sensitive to implied motion and more generally to stimuli that signal the actions of another individual (Hoffman & Haxby, 2000; Haxby, et al., 1999; Kanwisher, et al., 1997).

Hoffman and Haxby (2000) for the first time found that there is a discrepancy between the role of the FFA and the posterior STS (pSTS) (only the posterior part of the STS is one of the core face-selective regions in Haxby and colleagues' (2000) model) in face processing. In their one-back task, the FFA responded more to identity versus gaze direction, while the pSTS was more sensitive to gaze information than identity. Andrews and Ewbank (2004) found that the right pSTS was sensitive to faces with different viewpoints, facial expressions and also gaze direction, but not to faces with different identities. However, from their study, we cannot tell the role of pSTS in processing of separate viewpoint, expression, or gaze information, because they used face stimuli that differed in all three dimensions. Engell and Haxby (2007) further showed that gaze-direction and facial expression are actually represented by dissociable but also overlapping regions of the right pSTS. Activation of the pSTS to faces is not correlated with the behavioral face detection or identification performance (Grill-Spector, et al., 2004).

The amygdala, located in the anterior part of the medial temporal lobe, is involved in the representation of biologically relevant signals pertinent to survival. A recent model proposes a more general role of the amygdala to response to broader biologically relevant stimuli (Sander, Grafman, & Zalla, 2003). The amygdala is activated to facial expressions of both negative and positive emotions (Adolphs et al., 1999; Blair, Morris, Frith, Perrett, & Dolan, 1999; Phan, Wager, Taylor, & Liberzon, 2002; Winston, O'Doherty, & Dolan, 2003), even under unconscious conditions (Williams, et al., 2004; Gur, et al., 2002; Habel, et al., 2007). It is also sensitive to the direction of eye gaze, which is important for indicating the mental states of others (George, Driver, & Dolan, 2001; Kawashima et al., 1999; Richeson et al., 2003). To be more specific, the amygdala may be especially activated by increasing vigilance and attention based on stimuli ambiguity. And this process relies on the cognitive evaluation of meaning and outcome of the event, within a specific context and also relationship with personal goals (Sander et al., 2003). From this point of view, certain emotional faces (e.g., anger, fearful faces) represent biological information because they potentially conflict with one's goal and signal the presence of a danger for them. Studies also suggested that there are hemispheric differences in the representation of the amygdala to emotional stimuli, that is the right amygdala responds to rapid detection of emotional stimuli, and the left amygdala is involved in the more detailed emotional evaluations (Glascher & Adolphs, 2003; Wright et al., 2003).

Even though both the STS and the amygdala are responsive to facial expressions, one study has shown that the role of these two expression-selective regions is dissociated (Harris, Young, & Andrews, 2012). Facial expressions can be represented in both continuous and categorical ways in the brain. The right pSTS is sensitive to all changes in facial expressions, representing a continuous processing, whereas the amygdala is only sensitive to categories of expressions, indicating a categorical representation. A recent study by Flack and colleagues (2015) also showed that the pSTS does not encode facial expressions holistically, instead it reflects an early stage in the processing of facial expressions where facial features are represented separately.

Haxby and colleagues (2000) proposed a neurological model of face perception, which was compatible with the general ideas of the functional model of face perception by Bruce and Young (1986). The neurological model (see Figure 1.3) is comprised of a core system and an extended system. In the core system, the representation of invariant aspects of faces underlies recognition of face identity, whereas the representation of changeable aspects of faces, like facial expression, eye gaze, and lip movement, underlies perception of information that facilitates social communication. Within the core system, the OFA is involved in the structural-encoding phase of face processing, and then inputs information to both the FFA and the pSTS regions. A route to the FFA is involved in the high-level, identity-based encoding of faces. Another parallel route to the pSTS is involved in processing the changeable aspects of faces, such as facial expression and eye gaze direction. The pSTS also has connections with an extended network that includes the amygdala, which is involved in further evaluation of emotions.

The extended system includes regions that act in concert with the regions in the core system to process faces. To be more specific, it does not reflect the recognition of faces per se, but rather the associated and further processing of face stimuli. This system comprises processing of spatially directed attention (e.g. perception of gaze direction), prelexical speech perception (e.g. lip reading), emotion, and also personal semantic information (e.g. name, biographical informaiton).

This neural model also emphasizes a disassociation between the representations of invariant and changeable (such as eye gaze, expression, and lip movement) aspects of faces (Fox, et al., 2009; Ganel, Valyear, Goshen-Cottstein, & Goodale, 2005; Harris, Young, & Andrews, 2014b; Winston, Henson, Fine-Goulden, & Dolan, 2004). For example, repeating facial expressions reduced the fMRI signals in the FFA and the STS, while repeating facial identities only reduced signal in the FFA (Harris et al., 2012).

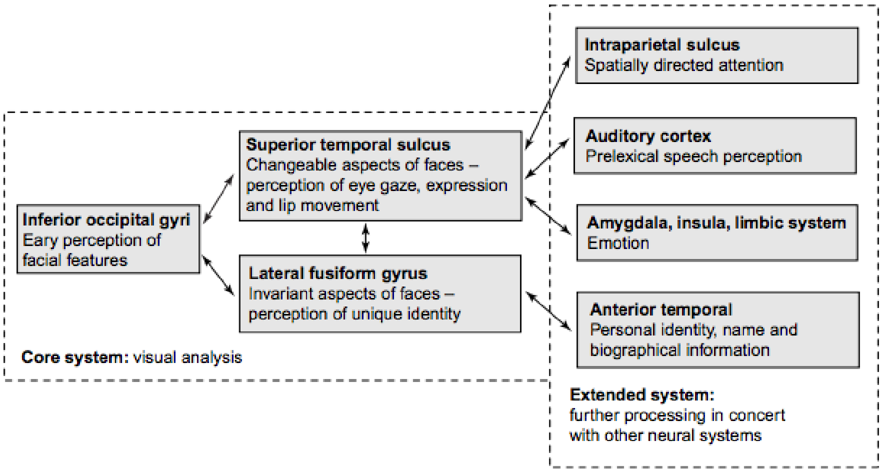


Figure 1.3 The distributed human neural system model for face perception (Haxby et al., 2000)

According to the neural model of face perception (Haxby et al., 2000), response differences to own-race and other-race faces could happen in the OFA, where physical changes between faces of different races are captured. Differences in activations could also happen in the FFA that is sensitive to facial identity, and also in the STS that is sensitive to facial expression. A summary of relevant studies can be found in the next section 1.4.3.

1.4.3 Neural mechanisms underlying the own-group advantage / other-race effect

Understanding the neural mechanisms that underlie how people process own-race versus other-race faces can help us better understand the nature of the other-race effect. The majority of the studies reported have been focused on examining differences between brain responses to White and Black faces in the United States, for largely historical reasons. However, differences in brain activations to Black versus White faces cannot be the whole story of the underlying neural mechanisms of the other-race effect; further studies still need to investigate the cultural differences in processing of faces across various races, such as the Western Caucasian and East Asian faces. Another important question is that most of the studies have been focused on exploring the mechanisms underlying the other-race effect on facial identity recognition; while the underlying mechanisms in facial expression processing still remain to be explored.

1.4.3.1 Cultural differences in activations of faces in the fusiform face area (FFA)

Studies have found that there are significant differences in FFA activations to own- versus other-race faces. In a classic recognition memory task, Golby et al. (2001) found that the FFA responded more strongly to own-race faces versus other-race faces for both European-American and African-American participants, but the memory differences between own- and other-race faces were only positively correlated with activation in the left FFA. Moreover, the greater activation in the FFA to own- versus other-race faces only results from observation of unfamiliar faces, rather than familiar faces (Kim et al., 2006). Using multivoxel pattern analysis (MVPA), Brosch and colleagues (2013) successfully predicted the race of faces from the BOLD activation patterns in the FFA in participants with high pro-White bias, indicating that strong implicit pro-White bias decreased the similarity of neural activations of White and Black faces in the FFA. However, Natu, Raboy and O'Toole (2011) only found a significant own-race bias in the broader ventral temporal cortex, rather than the FFA alone, but further analyses found the cross-cultural differences in the first few time points of the block for both the FFA and the broader ventral temporal cortex. Even faces from an arbitrary social category made by experimental manipulation (members of an experimentally created own-group) could create greater activations in the FFA than faces from a less relevant social group (members of an experimentally created out-group) (Van Bavel, Packer, & Cunningham, 2011).

1.4.3.2 Cultural differences in activations of faces in the amygdala

In contrast to the higher activation to own-race faces in the FFA, studies have shown greater amygdala BOLD activations to other-race faces than own-race faces, possibly reflected in the sensitivity to people's implicit attitude to own-race versus other-race members.

Hart et al. (2000) first used fMRI to examine the effects of race on the activation of the amygdala. They found no group differences in amygdala activation between own-race versus other-race faces during early presentations of faces, whereas activations in the amygdala declined significantly for own-race faces during later presentations. Following Hart et al. 's (2000) work, Phelps et al. (2000) found that differences in amygdala activation to Black versus White faces in White participants significantly correlated with their Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) scores, but not with the performance of an explicit measure of racial attitudes (the Modern Racism Scale; McConahay, 1986). For White participants with more negative implicit bias towards Blacks, there were greater amygdala activation responses to Black than White faces. The work of Phelps et al. (2000) indicates that implicit social attitudes to own-race versus other-race members play an important role in the neural activations in the amygdala. In addition, Wheeler and Fiske (2005) also found that activation of the amygdala to White and Black faces was modulated by social categorization. Black faces only elicited greater amygdala activation compared to White faces when participants' task was to socially categorize the faces. In addition, skin tone has also been found to moderate race-related amygdala activity; faces with Black skin tone can trigger greater amygdala activity than faces with White skin tone (Ronquillo et al, 2007).

As the amygdala is highly specialized for evaluating and responding to threatening signals, including fearful facial expressions, Chiao and colleagues (2008) investigated the neural mechanisms underlying the cultural differences in the recognition of fear expressions in the amygdala. Their results showed that both their Japanese and Western Caucasian participants showed greater amygdala activation to fear expressions of their own cultural group members. This is the only study that has shown the own-group advantages in facial expression processing.

1.4.3.3 Cultural differences in activation of faces in the anterior cingulate cortex (ACC) and the dorsolateral prefrontal cortex (dlPFC)

The anterior cingulate cortex (ACC) and the dorsolateral prefrontal cortex (dlPFC) are another brain areas activated in the race context. The ACC is involved in monitoring for response competition and may serve to engage executive control once a conflict is detected. Conflict between automatic, prepotent feelings and conscious control to respond fairly may explain the activation of ACC (Kubota, Banaji, & Phelps, 2012), whereas the dfPFC is involved in top-down executive control (Curtis & D'Esposito, 2003). Activations in the ACC are often found in combination with activations in the dlPFC in the majority of race studies.

Cunningham et al. (2004) investigated the neural mechanisms underlying conscious control towards Black versus White faces. Their results showed that Black faces triggered greater amygdala activation compared with White faces in White participants only when the faces were presented subconsciously. When presented consciously, greater activations in the dorsolateral prefrontal and the anterior cingulate areas were detected to Black compared to White faces. Increased frontal activation to Black faces also positively correlated with decreased amygdala activity to Black compared with White faces. In another study, Richeson et al. (2003) elucidated the roles of the dlPFC and the ACC in regulating the activation of race-related implicit attitudes. Their results indicated that the dlPFC is involved in engaging cognitive control during interracial contacts and the ACC is involved in detecting the need for engaging cognitive control.

1.5 The aim of this thesis

The overall purpose of this thesis is to investigate cultural similarities and differences in facial identity and expression processing and the underlying neural mechanisms in Western Caucasian and Chinese participants. As mentioned above, most studies investigating cultural differences in face processing have mainly focused on White and Black faces. Hence we know much less about the cultural differences in perception and recognition of faces of other races, for example the differences between Western Caucasian and East Asian faces. The present study will therefore bridge this gap by examining cultural differences between Western Caucasian and Chinese participants' processing of own- and other-race faces. In addition, most previous studies examining the other-race effect only focused on one racial group of participants or only one racial set of stimuli. This unbalanced design cannot offer a strong test of the other-race effect because recognition differences between the two sets of stimuli or two cultural groups might also be explained by other factors, such as the systematic differences between sets of stimuli, or the differences in decoding rules, instead of the other-race effect. In this thesis, the other-race effect will always be investigated with a fully balanced crossover design by testing Western Caucasian and East Asian participants with their own- and other-race stimuli. In addition, compared with many studies examining the cross-cultural differences in facial identity processing, this study will have a systematic investigation of the own-group advantage in facial expression, and also investigate the cross-cultural processing in these two domains simultaneously.

The main aims of the thesis are: (1) to investigate cultural similarities and differences in facial expression processing; (2) to examine the potential factors to explain the other-race effect in facial expression processing; (3) to look at the other-race effect for facial expression and identity processing at the same time and also to explore the extent of the cross-cultural differences and similarities; (4) to investigate the neural mechanisms of the cross-cultural processing of facial identity and expression.

Chapter 2 investigates the cultural similarities and differences in the perception and categorization of facial expressions of basic emotions. Also, following up recent studies suggesting that people from Chinese and British cultures use different information from faces to represent facial expressions (Jack et al., 2009; 2012), differences in information using from different part of faces (upper/lower regions) by people from East Asian (Chinese) and Western Caucasian (British) cultures are also explored by only presenting either the upper or lower half of faces.

Chapter 3 further explores the potential factors that drive cultural differences in expression recognition: (1) the confusions between certain expression pairs like anger and disgust or fear and surprise; (2) the differences in the engagement of holistic processing to expressions of different races. The composite-expression paradigm devised by Calder et al. (2000) is used to investigate the holistic processing of facial expressions. It is predicted that there will be differences in the engagement of holistic processing to own-race and other-race faces, for the reason that a number of studies have linked the own-group advantages in the recognition of facial identity to holistic processing (Tanaka et al., 2004; Michel et al., 2006a; Michel et al., 2006b). However, according to Jack et al.'s (2012) suggestion that East Asian participants mainly use information from the eye region to recognize facial expressions, holistic processing of expression should be reduced in comparison to Western Caucasian participants.

Chapter 4 aims to explore the cross-cultural difference in facial expression and identity processing at the same time with structure-matched tasks and also to investigate the magnitude of cross-cultural similarities and differences across these two domains. Even though studies have found reliable other-race effects in both facial expression and identity processing, no studies have yet examined these effects simultaneously. To date, substantial procedural differences between the tasks have precluded such a comparison, and the widely used methods also have significant limitations. Here in this chapter, a free photo-sorting tasked adapted from the task devised by Jenkins et al. (2011) is adapted to ask participants to sort photos into piles according to facial expression or identity. Using this task, this chapter can also investigate the magnitude of cultural similarities and differences in identity and expression processing by comparing participants' response patterns of photo-sorting.

Chapter 5 investigates the neural mechanisms in cultural processing of facial identity and expression within functionally defined regions of interest (ROIs) that are most sensitive to faces versus other objects, such as the FFA and the OFA. The fMR-adaptation method is used by showing participants own- and other-race facial images with either the same or different combinations of facial identity and expression: (1) same identity same expression;

(2) same identity different expression; (3) different identity same expression; (4) different identity different expression. The univariate ROI analysis is used to offer a better understanding of the neural representations of facial identities and expressions of different races.

2. Cultural Similarities and Differences in Perceiving and Recognizing Facial Expressions of Basic Emotions (Published in the Journal of Experimental Psychology: Human Performance and Perception, 42(3), 423–440)

2.1 Introduction

Facial expressions of emotion carry important social signals in daily communication. With increasing cross-cultural interaction and cooperation, understanding whether the processing of facial expressions is universal or culturally variable is a topic of both theoretical and practical importance. For much of the twentieth century, theories of cultural relativism held sway, and it was thought that facial expressions were primarily culturally constructed and learnt (see Elfenbein & Ambady, 2002b). However, for the last 40 years, the dominant position has been based on Ekman's (1980) interpretation of Darwin's (1872) proposal that a small number of basic emotions serve evolved biological functions, and that facial expressions of these basic emotions will be universal across cultures.

Many studies have provided evidence to support the universality hypothesis. For example, people can identify facial expressions of basic emotions portrayed by members of different cultures at above chance levels (Biehl, et al., 1997; Ekman, 1972; Izard, 1971), and people in preliterate cultures pose facial expressions that are similar to expressions used by people from Western cultures (Ekman, 1972). However, the correct interpretation of such findings is still debated. For example, the extent to which cross-cultural agreement is overestimated due to procedural factors such as the use of forced-choice recognition tests has been controversial (Ekman, 1994; Russell, 1994). In fact, even with forced-choice recognition, participants are more accurate at recognizing emotions expressed by members of their own cultural group (Izard, 1971; Ekman, 1972; Ekman, Sorensen, & Friesen, 1969). Studies have also revealed that this 'own-group advantage' is lower for groups with a closer geographical distance or having more cultural contact with each other (Elfenbein & Ambady, 2002a, 2002b, 2003a, 2003b, 2003c).

A number of recent studies have suggested that perception of facial expression is influenced by culture. For example, Yuki and colleagues (2007) found that East Asian participants give more weight to the eyes of either emoticons or face images compared to Western participants, whereas Western participants give more weight to the mouth region of the face when rating expressions in comparison with East Asian participants. Other studies have used eye movements and reverse correlation methods to determine the key features used to recognize facial expressions by Western Caucasian and East Asian participants. In an eye-tracking study, Jack et al. (2009) found that when recognizing facial expressions Western participants fixated features in the eye and mouth regions, whereas East Asian participants mainly fixated the eye region. Because eye fixation patterns indicate increased interest in certain regions of the face but do not rule out the possibility that features from less fixated regions are none the less encoded, Jack and colleagues (2012) used a reverse correlation method to determine whether these differences in fixation patterns reflect different mental representations of key expressive features. They asked participants to make forced-choice identification of the facial expressions of 12,000 highly ambiguous stimuli that were derived by adding pixel-based white noise to a neutral face base image. They then averaged the white noise templates associated with each categorical judgement by a participant, to try to capture that participant's internal representation of facial features of each facial expression. In line with their eye movement findings, Jack et al.'s (2012) results showed that Western participants used information from the eye and mouth regions to represent facial expressions internally, whereas East Asian participants relied largely on the eye region, including the eyebrows and eye gaze direction.

These studies can be related to a broader background of putative cultural differences between Western and East Asian participants, including claims that East Asian participants group objects 'based on family resemblance rather than category membership' (Nisbett & Masuda, 2003), and reports of cultural differences in perceptual fixation patterns even to non-emotional faces (Blais et al., 2008). Based on their findings, Jack et al. (2009, 2012) make a strong case for differences between Western and East Asian participants' mental representations of facial expressions, and specifically highlight differences in the use of the eye and mouth regions. However, there are some aspects of the techniques Jack et al. (2012) used that suggest further confirmation is needed. In particular, even though reverse correlation can in theory potentially capture any of the communicative signals participants might seek, in practice the potential variety of facial expression cues means that stimuli created by adding pixelated noise to a neutral face are unlikely to contain sufficient examples of all possible facial signals (Freeman & Ziemba, 2011). Hence, Jack et al.'s (2012) findings concerning how culture can finely shape the internal representation of facial expressions might be influenced by these constraints of their chosen reverse correlation method.

In the present study, we therefore followed up Jack et al.'s (2009, 2012) findings with different methods. In particular, we tested whether cultural differences between Chinese (East Asian) and British (Western) participants reflect differences in the perception of facial expressions of basic emotions (anger, disgust, fear, happiness, and sadness), or differences in the way that these expressions are categorized. To achieve this, two main tasks were used: (1) a perceptual similarity task, and (2) a categorization task. The perceptual task involved rating the degree of similarity in expression between pictures of facial expressions of same or different emotions posed by different individuals. In this task, the pairs of images to be rated were always very different in themselves (because they showed different individuals), but raters were asked to ignore the differences in identity and focus only on how similar or different the facial expressions were. This task was used to generate a matrix of perceived similarities between exemplars of facial expressions of the five basic emotions for Chinese and for Western participants, and is equivalent to the kind of analysis used to create well-known perceptual models such as Russell's circumplex (Russell, 1980). The categorization task involved forced-choice recognition of emotion from the same images as were used in the perceptual similarity task. This task was used to compare recognition rates between Chinese and British participants. To achieve a systematic evaluation of the causes of cultural differences in perception and recognition, we used stimuli drawn from three different sets of expressions posed by Western and by Chinese participants and tested responses based on the whole face, the upper (eyes and forehead), or the lower (mouth and chin) part of each face.

2.2 Experiment 1

This experiment aimed to explore the cultural differences in the perception and recognition of facial expressions of basic emotions between Chinese and British participants with the most widely-used set of facial expressions - the Ekman and Friesen (1976) series. This set has been used in hundreds of studies because the expressions are well-validated and based on a careful analysis of underlying muscle movements that can create a plausible apex expression for each emotion (Bruce & Young, 2012). We used examples from the Ekman and Friesen (1976) set selected from the FEEST series (Young et al., 2002). The basic emotions represented were anger, disgust, fear, happiness and sadness. Although also present in FEEST, expressions of surprise were omitted because the status of surprise as a basic emotion has been questioned (Oatley & Johnson-Laird, 1987); one can be pleasantly or unpleasantly surprised (Du, Tao, & Martinez, 2014).

2.2.1 Methods

2.2.1.1 Participants

Three different levels of difficulty of the perceptual similarity task were used in this experiment, while the categorization task was always the same. Three groups of participants from the University of York were therefore recruited separately for the different levels of difficulty of the perceptual similarity task: (1) 20 Chinese students (mean age, 21.7 years old) and 20 British students (mean age, 19.5 years old); (2) 9 Chinese students (mean age, 21.6 years old) and 9 British students (mean age, 19 years old); (3) 10 Chinese students (mean age, 19.9 years old) and 10 British students (mean age, 18.9 years old). In addition to those reported above, data for a further 4 participants (2 Chinese, 2 British) were excluded because of failure to comply with the instructions (they rated all pairs of expressions as equally different). Based on self-report all British participants were of white Caucasian ethnic background, and Chinese participants were brought up in China with Chinese parents. All participants gave their consent prior to the experiment and received a small payment or course credit. The University of York Department of Psychology Ethics Committee approved the study.

2.2.1.2 Stimuli

We used the same set of Ekman and Friesen faces selected from the FEEST set as used in recent studies (Mattavelli, et al., 2013; Harris, Andrews, & Young, 2012, 2014a, 2014b). This set of stimuli comprises photographs of five individuals each posing facial expressions of five basic emotions (anger, disgust, fear, happiness, and sadness). The images were selected based on the following three main criteria: (1) a high recognition rate for all expressions (mean recognition rate in a six-alternative forced choice experiment: 93%), (2) consistency of the action units (muscle groups) across different individuals posing a particular expression, and (3) visual similarity of the posed expression across individuals. Using these criteria to select the individuals from the FEEST set helped to minimize variations in how each expression was posed. Ten additional facial images with two actors posing five expressions were randomly chosen from the FEEST set to use for the practice run. The resolution of each face image was 420 pixels high and 280 pixels wide. When viewed from 57 cm away, each image extended 11.1 degrees high and 7.5 degrees wide.

2.2.1.3 Procedure

Participants viewed images of facial expressions using a computerized task programmed with PsychoPy software (www.psychopy.org). All the participants had to complete the perceptual similarity task first, and then the expression categorization task. In the perceptual similarity task, participants saw two facial expressions posed by different actors and their task was to rate the similarity of these two facial expressions on a 7-point scale, with 1 indicating not very similar expressions and 7 very similar expressions. The two facial expressions could represent the same or different emotions, but the expressions in each pair were always posed by different actors and the participants were asked to avoid rating the similarity of facial identity across the two faces and focus on their expressions. Participants were asked to rate the similarity of the facial expressions in fifteen different types of expression pairs (same-expression pairs: anger-anger, disgust-disgust, fear-fear, happiness-happiness, sadness-sadness; between-expression pairs: anger-disgust, anger-fear, anger-happiness, anger-sadness, disgust-fear, disgust-happiness, disgust-sadness, fear-happiness, fear-sadness, and happiness-sadness). Because each emotional expression was posed by five actors and the two expressions presented for rating in any one trial were always posed by different actors, there were a total of ten possible combinations for each of the five same-expression pairs, leading to a total of 50 same expression pair trials. Similarly, there were twenty possible combinations for each of the ten between-expression pairs. Therefore, each participant had to complete a total of 250 trials in the rating task. These were presented in random order, with a short break permitted following the completion of the first 125 expression pairs. Ten additional trials were included to form a practice run at the start of the experiment.

Three different variants of this perceptual similarity rating task were used because we wanted to examine whether task factors involving overall difficulty levels and the degree of emphasis on initial perceptual encoding might affect the degree of apparent similarity across cultures in perceiving facial expressions. The differences between the three perceptual tasks were as follows:

(a) Perceptual similarity task with simultaneous presentation of two face stimuli. In this task, the two faces were presented simultaneously side by side next to the middle of the screen for 5 seconds, allowing time for encoding and comparing the images while both were visible. The rating scale remained on the screen until the participant made a response.

(b) Perceptual similarity task with sequential presentation of two face stimuli. The two faces were presented successively in this task, separated by a 2-second fixation interval. The first face image was presented for 1.5 seconds, and the second one for 2 seconds. This task was intended to be more difficult than task (a) because participants had to cross-refer their encoding of the second face to their memory of the first face before they could make a similarity rating.

(c) Perceptual similarity task with sequential presentation and an interpolated mask (see example in Figure 2.1). In this task, the two face stimuli were presented sequentially, separated by a phase scrambled face image mask for 1 second and also a 1-second fixation screen. The first face image was presented for 1.5 seconds, and the second one for 2 seconds. In each trial, the first facial image was always followed by a facial mask that was derived by phase-scrambling the neutral facial image expressed by the same actor. This task was even more difficult than task (b) because the facial mask that followed the first face would interrupt participants' visual representation of the facial expression of the first face, making it harder to make a comparison with the second face.

In the Categorization task (see Figure 2.1), only one face image was presented on the screen for each trial, and participants were required to perform a five-alternative forced-choice (5AFC) task to identify its facial expression as anger, disgust, fear, happy or sad. Each face image remained visible on the screen until the participant made a response. The sequence of the emotion labels was consistent across all participants. The 25 face images (5 models posing 5 basic emotions) were presented twice each for each participant in a random order, yielding a total of 50 trials. They were also given a practice task of identifying 10 other facial expression pictures before the formal experiment.

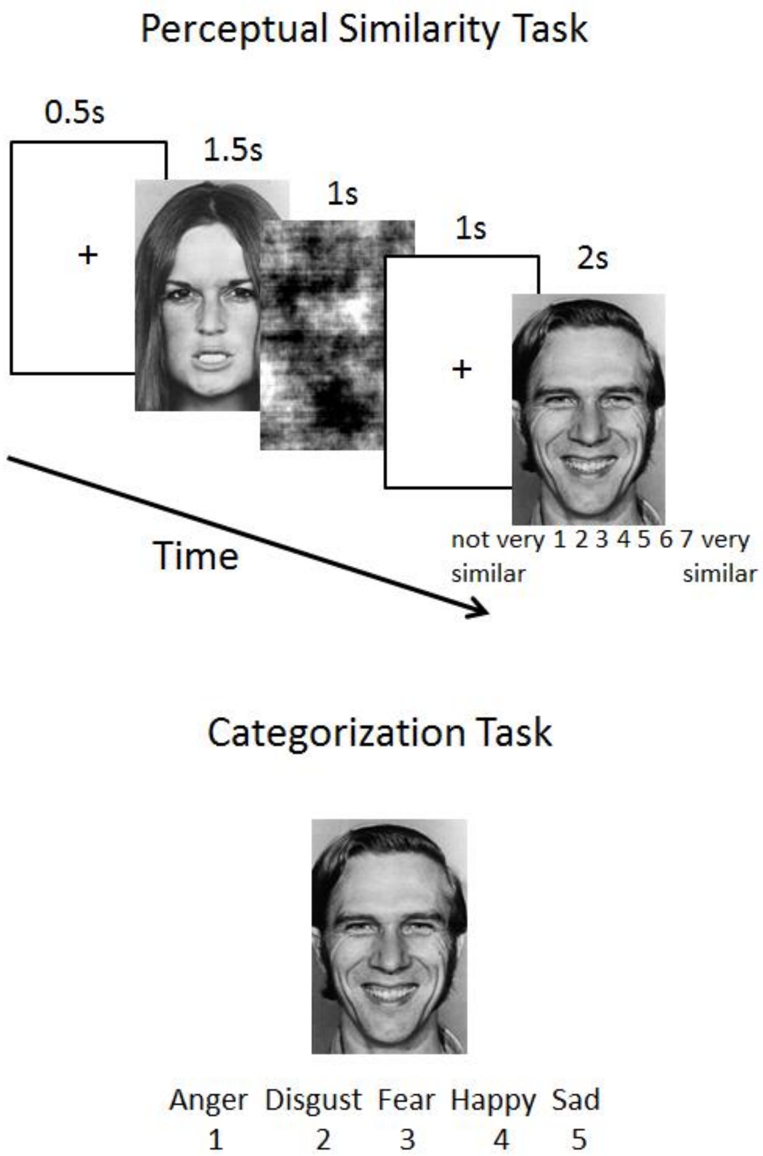


Figure 2.1 Examples of stimuli for the two tasks: Perceptual similarity task (version c) and Categorization task in Experiment 1. All face images are from the Ekman and Friesen (1976) set selected from the FEEST series (Young et al., 2002).

Instructions for the Western participants were given in English. Instructions for Chinese participants were translated into Chinese by the experimenter, but the five emotion labels shown on screen in the categorization task remained in English, and Chinese participants were asked to write down the Chinese meaning of these five emotion words immediately after they finished the categorization task (see Appendix 1). The English emotion labels used in the categorization task were correctly understood by all Chinese participants, and this was verified by three native Chinese speakers. After completing the experiment, all Chinese participants completed a questionnaire about how long they had been in the UK (see Appendix 2).

2.2.2 Results

To analyse the data from each variant of the perceptual similarity task, we calculated the average similarity rating for each pair of expressions for each participant (anger-anger, anger-disgust, anger-fear, etc). The resulting fifteen averaged ratings were then used to create a matrix reflecting the perceived similarity between expressions for participants from each cultural background (see Figure 2.2). This then allowed us to measure the overall concordance between the ratings of British and Chinese participants by correlating the obtained values for rated perceptual similarities across each group. The representational similarity matrices for Chinese and British participants and the correlations between the two groups with each version of the perceptual task are shown in Figure 2.2. These correlations were very high for each version of the task, $r = 0.99$, $p < .001$ for variant (a), $r = 0.99$, $p < .001$ for variant (b), and $r = 0.99$, $p < .001$ for variant (c), indicating that the overall pattern of perception of the expressions between the two cultural groups was strikingly similar across all levels of task difficulty.

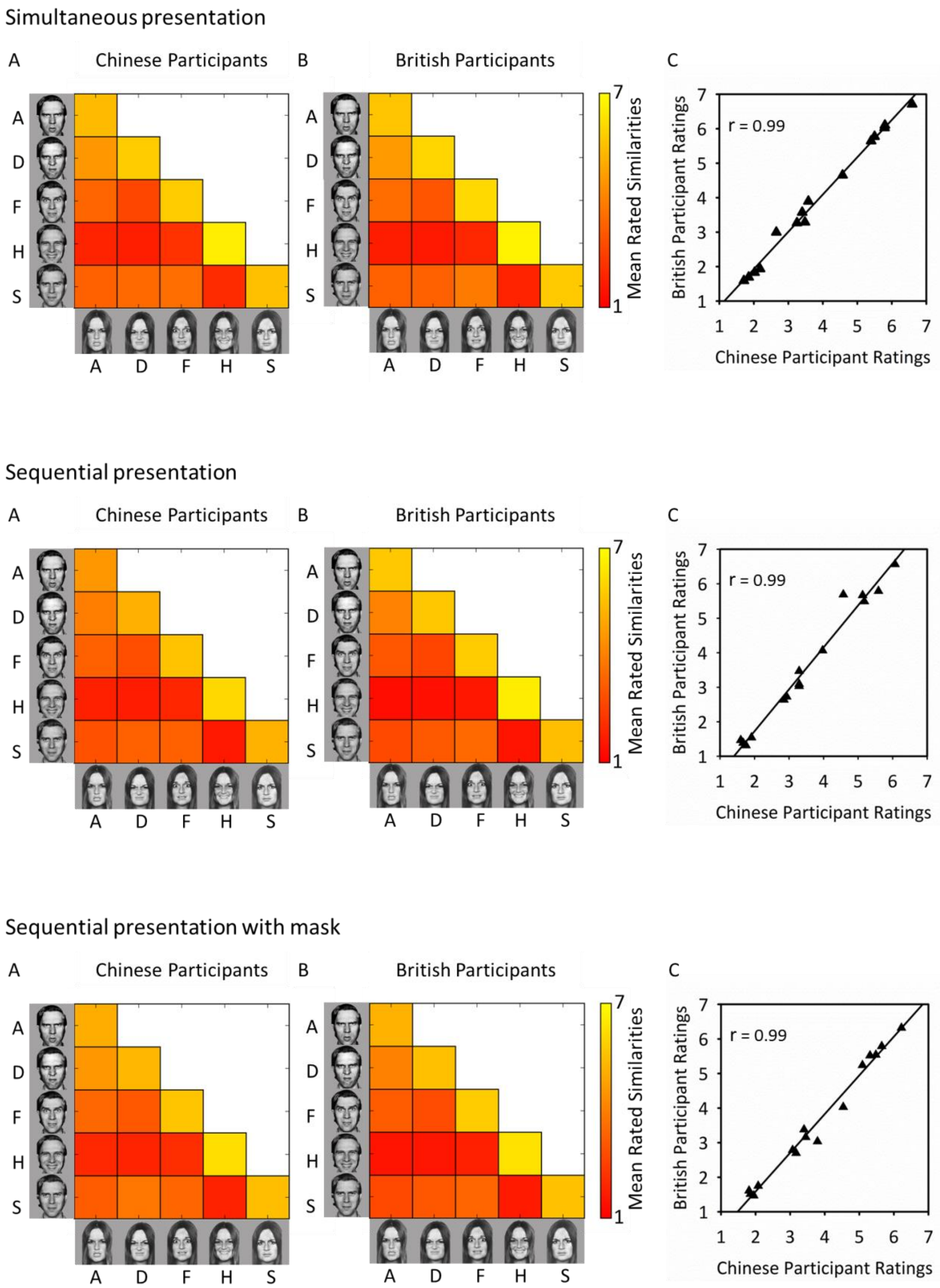


Figure 2.2 Correlation analyses of similarity ratings for three different versions of the perceptual similarity task (version a - top row, simultaneous presentation) (version b - middle row, sequential presentation) (version c - bottom row, sequential presentation with an intervening mask) between Chinese and British participants in Experiment 1. Similarity matrices for (A) Chinese and (B) British participants (A: anger, D: disgust, F: fear, H: happy, S: sad). (C) Scatterplot of rating correlation between the two groups of participants. All face images are from the Ekman and Friesen (1976) set selected from the FEEST series (Young et al., 2002).

Although the above procedure is sufficient to establish a close overall concordance between Chinese and Western perception of expression, it is in principle possible that some of this concordance might be driven by the relatively high similarity ratings for same-expression pairs (anger with anger, disgust with disgust, etc) that fall along the long diagonals in the representational similarity matrices in Figure 2.2. We therefore recalculated the correlations with ratings of these same-expression pairs removed, leaving only ratings of the ten combinations of different-expression pairs. In this way, we were able to estimate the structure of between-category differences themselves (for example, whether expressions of anger are perceived as more like disgust than happiness). Again, strikingly high correlations between the ratings of Chinese and British participants were obtained; $r = 0.98$, $p < .001$ for variant (a), $r = 0.99$, $p < .001$ for variant (b), and $r = 0.98$, *p*

< .001 for variant (c). The perceptual rating task therefore showed near-identical patterns across Chinese and British participants, regardless of task variations.

All 39 Chinese and 39 British participants did the same categorization task, allowing an overall analysis. A mixed-ANOVA was conducted on the arcsin transformed recognition accuracies, with Expression (anger, disgust, fear, happiness, sadness) as a within-subjects factor and Group (Chinese, British) as a between-subjects factor. Results showed a small but significant main effect of Group ($F(1,76) = 8.13$, $p < .01$, partial $\eta^2 = 0.10$), with British participants performing slightly better at categorizing these facial expressions than Chinese participants. The overall percentage recognition accuracies for each group are shown in Figure 2.3. There were also main effects of Expression ($F(4,304) = 38.18$, $p < .001$, partial $\eta^2 = 0.33$) and a significant interaction between Expression and Group ($F(4,304) = 3.40$, $p = .01$, partial $\eta^2 = 0.04$). Further analysis of this interaction showed that British participants were slightly better at identifying anger ($t(76) = 2.70$, $p < .01$) and disgust ($t(76) = 3.36$, $p < .001$) expressions than Chinese participants.

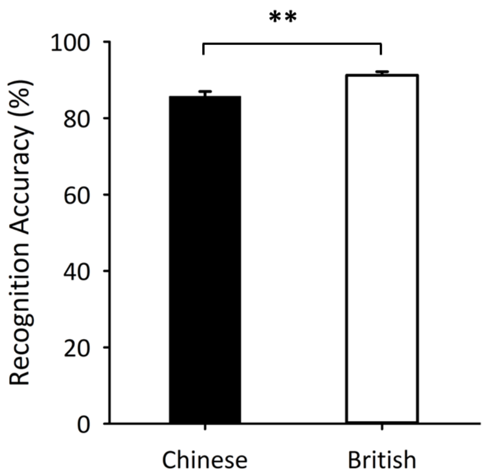


Figure 2.3 Overall emotion categorization accuracies (with standard error bars) for Chinese and British participants in Experiment 1. Asterisks denote higher overall emotion recognition rate for British participants than Chinese participants. **: $p < .01$.

Although our principal focus of interest was in the patterns of perceptual similarity and the accuracy of categorization of facial expressions across cultures, we were also able to check whether there were differences in response times. To do this, we analysed the reaction times (RTs) for both the perceptual similarity task and the categorization task. In the perceptual similarity task, a one-way ANOVA with Group (Chinese or British participants) as the independent variable and RTs as the dependent variable was conducted. Our results did not show an effect of Group, $F(1,75) = 0.93$, $p > .1$, partial $\eta^2 = 0.12$. In the categorization task, trials with incorrect responses were excluded and a one-way ANOVA with Group as the independent variable on the medians of correct RTs was conducted. Again, we did not find a significant Group effect, $F(1,75) = 0.93$, $p > .1$, partial $\eta^2 = 0.12$. These results indicate that there were no time differences for the participants in perceiving and identifying expressions of their own or the other-race group. They also confirm that the results in the categorization task did not result from any speed-accuracy trade-off.

From the Chinese participants' responses to the questionnaire, the time period that they had been in the UK varied from one month to nine years, but 27 out of 39 participants had been living in the UK for or less than a year. In order to investigate whether or not the amount of time that Chinese participants had lived in a western environment might affect their perception of facial expressions, we took the average similarity ratings of each of the 39 participants in the Chinese group and correlated these with the average ratings of the matched set of British participants on the equivalent variant of the perceptual task. These correlations with overall British performance for the 39 Chinese participants were then correlated with their time spent in the UK, as shown in Figure 2.4A. The overall correlation was non-significant, $r = 0.07$, $p = .66$.

A correlation analysis was also conducted to evaluate any relationship between each Chinese participant's recognition accuracy in the categorization task with their time in the UK. The results again did not find a significant correlation between categorization performance and time in the UK ($r = 0.15$, $p = 0.36$) (see Figure 2.4B). These results offer no evidence that the similarities and differences in perception and recognition of facial expressions were affected by the amount of time Chinese participants had spent in a western environment.

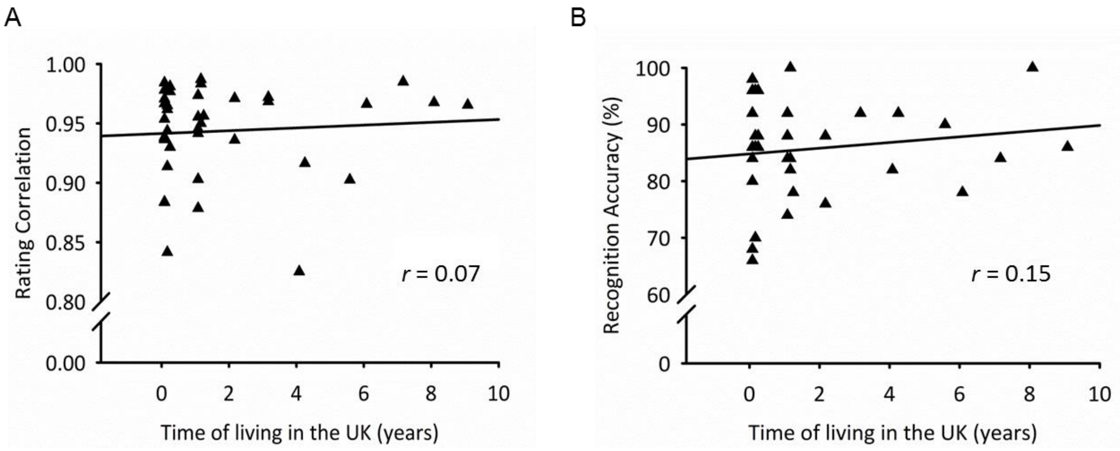


Figure 2.4 Scatterplots of Chinese participants' time living in the UK with their performance in the perceptual similarity task (A) and the categorization task (B) in Experiment 1.

2.2.3 Discussion

In this experiment, we investigated differences between Chinese and British participants in the perception and categorization of facial expressions from the Ekman and Friesen series. Our results revealed a potential divergence in the way people from these two cultures perceive and recognize facial expressions of basic emotions. In the perceptual similarity tasks, we did not find differences in the patterns of responses between Chinese and British participants, even in the most demanding version of the task that required the participants to remember the encoding of a masked facial expression. Instead, there was a high consistency in the rated similarity of expressions across cultures, showing that participants from Chinese and British cultures see facial expressions of basic emotions in much the same way at the perceptual level. Since we did not find group differences in the pattern of perceptual similarities between Chinese and British participants in even the most difficult task, we therefore only used the most difficult perceptual similarity task (successive presentation with a mask) in the following experiments.

In contrast to the perceptual similarity task, a small but statistically reliable cultural difference (5.23%) was found between Chinese and British participants in the categorization task. British participants were slightly better at categorizing facial expressions than Chinese participants. As we used images of western-looking individuals from the Ekman and Friesen series, this result is consistent with previous findings indicating the possible existence of an own-group advantage (Elfenbein & Ambady, 2002b; Biehl, et al., 1997; Ekman, 1972; Izard, 1971). The time that Chinese participants had been living in the UK did not have a significant effect on their perception or categorization of facial expressions.

The results of this experiment indicate that culture may slightly shape the way people from Chinese and British cultures recognize facial expressions, but not the perception of the facial expressions themselves. However, Jack and her colleagues (2012) suggested that East Asian and Western participants expect facial expressions to be primarily signalled from different regions of the face, making it possible that our initial tactic of using whole faces as stimuli may have reduced the impact of some of these differences. We therefore decided to investigate this possibility in Experiment 2.

2.3 Experiment 2

According to Jack et al. (2012), East Asian participants tend mainly to focus on the upper region of faces to internally represent facial expressions, whereas Western participants use the upper (eyes and eyebrows) and lower (mouth) regions more equally. Therefore, in this experiment, we presented only the upper or lower part of Ekman and Friesen faces to Chinese and British participants to further investigate potential strategy differences they might use in perceiving and categorizing facial expressions.

2.3.1 Methods

2.3.1.1 Participants

Twenty Chinese (mean age, 21.5 years old) and 20 British (mean age, 20.3 years old) participants were recruited to do the perceptual similarity rating and emotion categorization tasks with only the presentation of the upper half of each face, while another 20 Chinese (mean age, 21.6 years old) and 20 British (mean age, 19.3 years old) participants did the same two tasks viewing only the lower half faces. All participants were given a small a payment or course credit. In the lower region session, one Chinese participant's data was deleted mistakenly from the results, leaving only 19 participants in the Chinese group for analysis.

2.3.1.2 Stimuli

The images used in experiment 1 were again used in this experiment, except that they were divided into upper and lower half faces. Upper and lower halves of the faces were divided by a horizontal line through the middle of the bridge of the nose. The upper region faces were presented with a grey mask covering the lower part, and the lower region faces were presented with a grey mask covering the upper part (see Figure 2.5).

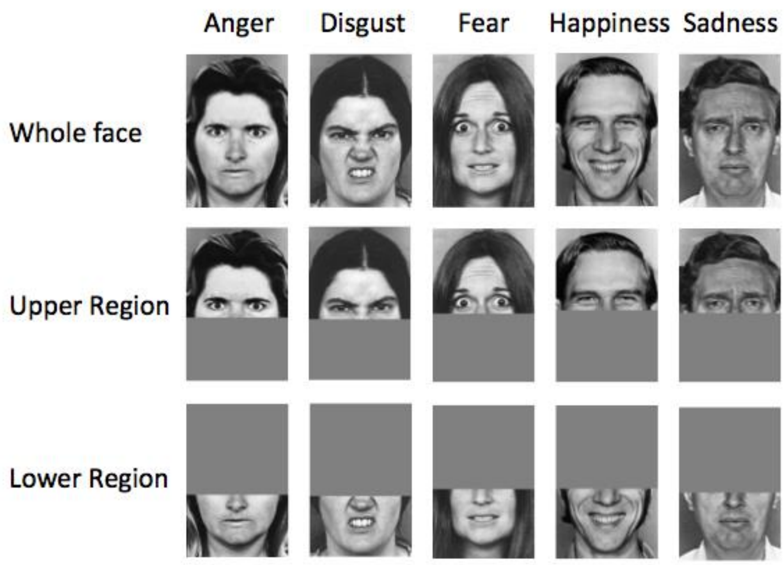


Figure 2.5 Examples of whole, upper and lower half region face images for each expression posed by a different model from the FEEST set. Only the upper and lower region face images were used in Experiment 2. All face images are from the Ekman and Friesen (1976) set selected from the FEEST series (Young et al., 2002).

2.3.1.3 Procedure

Participants were asked to carry out version (c) of the perceptual similarity task (successive presentation with a mask) and the categorization task in this experiment. Apart from the use of only upper half or lower half face images, all the other procedure details were identical to Experiment 1. As for Experiment 1, all Chinese participants showed correct understanding of the five emotion labels used in the categorization task.

2.3.2 Results

We conducted the same analyses as those used for Experiment 1. In the perceptual similarity task, the correlation between rated similarities across all expression pairs between Chinese and British participants was very high for the upper half faces, $r = 0.99$, $p < .001$, and for the lower half faces, $r = 0.99$, $p < .001$. The average ratings for each group of participants and scatter plots are shown in Figure 2.6.

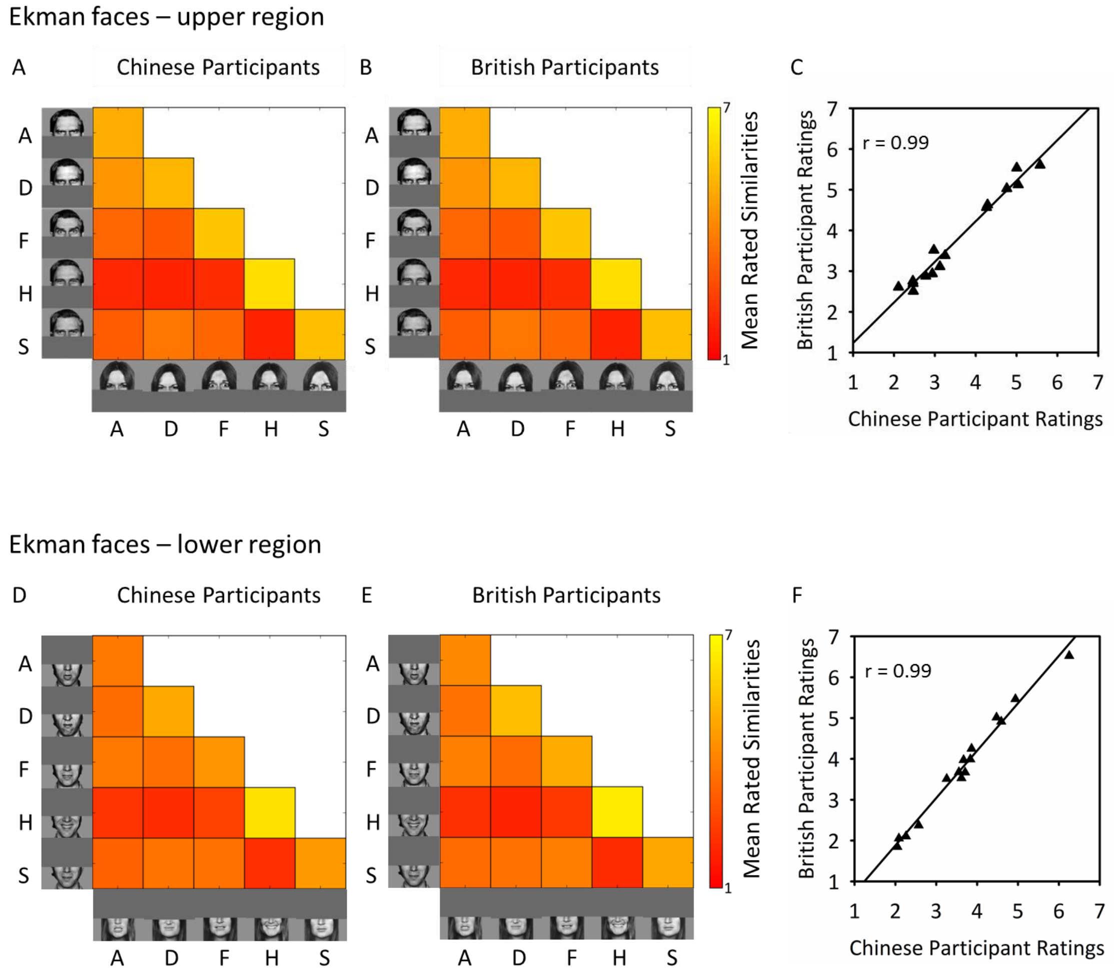


Figure 2.6 Correlation analyses of similarity ratings between Chinese and British participants in Experiment 2. Similarity matrices for (A) Chinese and (B) British participants for presentations of the upper face region (A: anger, D: disgust, F: fear, H: happy, S: sad). (C) Scatterplot of rating correlation between the two groups for the upper region. Similarity matrices for (D) Chinese and (E) British participants for presentations of the lower face region. (F) Scatterplot of rating correlation between the two groups for the lower region. All face images are from the Ekman and Friesen (1976) set selected from the FEEST series (Young et al., 2002).

These correlations were still high when the same-expression pairs (anger-anger, disgust-disgust, etc) were removed from the analyses; the correlation of between-category expression pairs was 0.95 ($p < .001$) for the upper face region, and 0.99 ($p < .001$) for the lower region.

In the categorization task, as would be expected, recognition rates for emotion from part faces were lower than the rates for whole faces reported in Experiment 1 (Whole face: 0.89 (0.15). In Experiment 2, mean recognition rates for emotion were as follows; Upper half face: 0.75 (0.27); Lower half face: 0.70 (0.24)). A 5 (Expression: anger, disgust, fear, happiness, sadness) × 2 (Group: Chinese, British) × 2 (Face Region: upper, lower) mixed-ANOVA on arcsin transformed recognition accuracies found a significant interaction of Group x Face Region ($F(1,75) = 9.08$, $p < .01$, partial $\eta^2 = 0.11$) (see Figure 2.7), which forms the principal focus of interest in the categorization data. Further analyses of this interaction revealed no group difference between Chinese and British participants for the upper half faces ($t(38) = 1.33$, $p > .1$), whereas for the lower region faces the British participants were slightly more accurate at categorizing the facial expressions than Chinese participants ($t(37) = 3.30$, $p < .01$). Note again that with the Ekman and Friesen faces the part expressions are posed by members of their own ethnic group for the British participants.

Main effects of Expression ($F(4,300) = 99.17$, $p < .001$, partial $\eta^2 = 0.57$) and Face Region ($F(1,75) = 7.52$, $p < .01$, partial $\eta^2 = 0.09$) were also found, and a significant two-way interaction of Expression × Face Region ($F(4,300) = 54.6$, $p < .001$, partial $\eta^2 = 0.42$). Further analyses showed that anger ($t(77) = 3.23$, $p < .01$), fear ($t(77) = 9.31$, $p < .001$), and sadness

($t(77) = 4.65$, $p < .001$) expressions were better recognized from the upper region of the faces than the lower part, whereas disgust ($t(77) = 8.23$, $p < .001$) and happiness expressions ($t(77) = 1.97$, $p = .05$) were better identified from the lower part of faces than the upper part. These findings are in line with previous results for recognition of emotion from parts of Ekman and Friesen faces (Calder, Young, Keane, & Dean, 2000). The three way interaction of Expression × Face Region × Group was borderline but not significant, $F(4,300) = 2.14$, $p = .08$, partial $\eta^2 = 0.03$. Other effects were not significant, either.

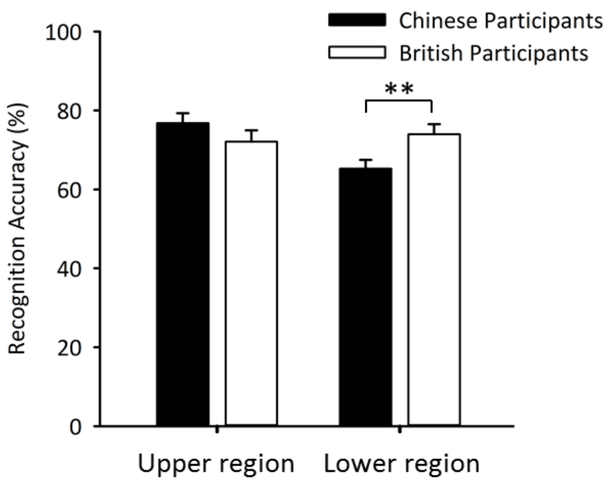


Figure 2.7 Overall emotion categorization accuracies (with standard error bars) for Chinese and British participants from upper and lower face regions in Experiment 2. Asterisks indicate the higher recognition rate for British participants in comparison with Chinese participants from the lower face region. **: $p < .01$.

Again, we analysed response times to check whether there were any group differences in RTs to own and other-race faces. For these RT analyses, we did not find any significant main effects or interactions with Group in both the perceptual similarity task and the categorization task. These results indicated that there were no time differences for the participants in perceiving and identifying expressions of their own or the other-race group, and that there was no speed-accuracy trade-off in the categorization task.

The time that Chinese participants had been in the UK varied from three months to six years, but was less than a year for 21 of 39 participants. To investigate the effect of the time of living in a western environment on participants' processing of facial expressions, we conducted the same correlation analyses as those used for Experiment 1 for both the upper and lower region faces. For the upper region of faces, the correlations between the time Chinese participants had lived in the UK and their performance in the perception and categorization tasks were -0.03 (p = .89) and -0.06 (p = .82), respectively. For the lower region session, the correlations between time in the UK and Chinese participants' performance in the perception and categorization tasks were -0.12 (p = .63) and -0.02 (p = .95), respectively. Overall, our analyses again found no significant effect of time spent living in the UK on Chinese participants' performance.

2.3.3 Discussion

In Experiment 2 we investigated potential cultural differences in perceiving and categorizing facial expressions by showing only the upper or lower half region of the faces to the participants. Again, we did not find any group differences in the patterns of perceptual similarity, with high correlations across Chinese and British participants for both the upper and lower face regions. Together with the results we found in Experiment 1, the first two experiments strongly indicate that people from Chinese and British cultures see similarities and differences between facial expressions from the Ekman and Friesen set in much the same way as each other. Moreover, they appear to use information from the upper part or from the lower half of the face in the same ways in the perception of facial expressions.

In the categorization task, even though recognition rates for half faces were lower than those for the whole faces in Experiment 1, none of the upper or lower regions of the expressions were recognized at anywhere near chance level (0.20). This result meant that, for all five facial expressions, both sections of the face contained emotional information that could be recognized by participants. None the less, the results also indicated a possible own-group advantage for the lower half faces. To be more specific, no reliable differences were found in recognition accuracy between the two cultural groups with the presentation of upper half faces, whereas British participants performed slightly better than Chinese participants at recognizing facial expressions in the lower face region.

The finding that British participants were better at categorizing expressions from the lower region of faces than Chinese participants is consistent with Jack et al.'s (2009, 2012) study, in which they found that Chinese participants mainly focused on the eye region of faces in categorizing facial expressions, whereas British participants focused more evenly on the eye and mouth regions. Our results were also consistent with the findings of Calder et al. (2000)

that anger, fear and sadness were better recognized from the upper region of the face, whereas disgust and happiness were better identified from the lower region of the face.

However, there are also some limitations to the data from the first two experiments. Firstly, we only used Western faces as stimuli. This unbalanced design does not offer a strong test of the own-group advantage because recognition differences between the two cultural groups might also be explained by differences in emotion decoding rules regardless of the stimuli being used (Matsumoto, 2002). Secondly, the stimuli we used portrayed highly standardized facial expressions that were created according to Ekman and Friesen's Facial Action Coding System (FACS, Ekman, & Friesen, 1978). This might also be limiting because expressions occurring in daily life might actually be more varied than those we used in Experiments 1 and 2. It is therefore important to establish whether or not the same pattern of results would also hold across fully balanced sets of Chinese and Western faces showing emotional expressions with the degree of variability we may encounter in everyday life.

2.4 Experiment 3

In Experiment 3, we used sets of faces portrayed by models from both Chinese and Western Caucasian cultures. In contrast to the Ekman and Friesen stimuli, which are based around prescribed muscle Action Units, the facial expressions in these two sets were developed by asking actors to pose facial expressions by imagining certain emotional scenarios. This way of eliciting stimuli leads to more varied expressions that may represent some of the diversity that exists in the natural world. The aim of Experiment 3 was further to explore the pattern of results from Experiment 1 by using varied expressions with both Chinese and Western Caucasian faces as stimuli.

2.4.1 Methods

2.4.1.1 Participants

Experiment 3 involved a larger set of stimuli than Experiments 1 and 2, so different groups of participants were recruited for the perceptual similarity task and the categorization task to minimise effects of task fatigue on the results. Twenty Chinese (mean age, 20.9 years old) and 20 British (mean age, 20.1 years old) participants took part in the perceptual similarity task, while another 20 Chinese (mean age, 23.4 years old) and 20 British (mean age, 21.1 years old) participants took part in the categorization task. All participants were given a small payment or course credit.

2.4.1.2 Stimuli

Two sets of images showing facial expressions of five basic emotions (anger, disgust, fear, happiness, and sadness) were selected from (1) the Chinese Facial Affective Picture System (CFAPS) (Wang & Luo, 2005; Gong, Huang, Wang, & Luo, 2011) posed by Chinese participants, and (2) the Karolinska Directed Emotional Faces (KDEF) (Lundqvist, Flykt, & Öhman, 1998) posed by Western Caucasian participants. Both sets of images were developed by asking actors to pose strong and clear facial expressions by imagining certain emotional scenarios.

There are in total 528 face images from 5 expression categories (anger, disgust, fear, happiness, and sadness) in the Chinese face set, and 700 expression images with full-face pose in the KDEF set. In order to explore potential own-culture and other-culture differences, it was important to establish that the faces were seen as being of 'Chinese' or 'Western' appearance. We therefore piloted 200 facial expressions from each set involving 40 randomly chosen examples of each of the 5 emotions, to identify expressions that were reliably seen as being posed by Chinese or Western models. These 200 Chinese and 200 KDEF faces were shown to an additional sample of 12 participants (6 Chinese and 6 British), asking them to decide whether each image was that of a Chinese or Western individual. From these data we selected final sets of 100 Chinese and 100 KDEF faces (with 20 exemplars of each of the 5 emotions) that were reliably seen as being of 'Chinese' or 'Western' appearance for use in the categorization task. The overall rates at which these were seen to represent Chinese or Western models were 99.6% and 98.6%, respectively.

In order to match the characteristics of the perceptual similarity task with those used in the first two experiments, 25 Chinese and 25 KDEF face images (5 exemplars for each expression category) were selected from each 100-image set used in the categorization task.

The full-face images chosen from the KDEF set were converted to greyscale and cropped, to match the general appearance of the faces in the Chinese set. The luminance values of all the KDEF faces were also adjusted to match the overall luminance of the Chinese faces. All face images were resized to 300 pixels high and 260 pixels wide, and when viewed from 57 cm away, each image extended approximately 8 degrees high and 7 degrees wide. Figure 2.8 shows examples of images used in the following two experiments.

Figure 2.8 Examples of whole, upper and lower half region face images for each expression posed by a different model from the Chinese and KDEF sets. The whole faces were used in Experiment 3, and the upper and lower region face images were used in Experiment 4. All Chinese face images are from the Chinese Facial Affective Picture System (CFAPS) (Wang & Luo, 2005; Gong et al., 2011) and all Western Caucasian faces are from the Karolinska Directed Emotional Faces (KDEF) (Lundqvist et al., 1998). The id of the KDEF images shown here are AM24ANS, AF09DIS, AF01AFS, AF08HAS, and AF26SAS, respectively.

2.4.1.3 Procedure

Participants were required to do either version (c) of the perceptual similarity task (successive presentation with a mask) or the categorization task with the presentation of the Chinese and KDEF faces. In the categorization task, the sets of 100 Chinese and 100 KDEF faces were each divided randomly into 2 blocks, yielding a total of 4 blocks of 50 faces.

Images were then presented to each participant in a block order of 'Chinese-KDEF-Chinese-KDEF' or 'KDEF-Chinese-KDEF-Chinese', which was counterbalanced across participants. Each image was presented for 1 second. The order of the emotion labels used for the categorization responses was also counterbalanced across participants. Apart from the face stimuli used and the changes in the categorization task resulting from the shorter presentation time and the use of blocked presentation of Chinese and KDEF faces mentioned above, all other procedure details were the same as those for Experiment 1. In addition, all Chinese participants showed correct understanding of the meaning of the five emotion labels used in the categorization task.

2.4.2 Results

The same analyses were conducted as those used for Experiment 1. In the perceptual similarity task, ratings of all expression pairs between Chinese and British participants showed a very high correlation for the Chinese faces, $r = 0.99$, $p < .001$, and also for the KDEF faces, $r = 0.98$, $p < .001$. The representational similarity matrices and scatter plots for the two groups of participants and the two sets of faces are shown in Figure 2.9.

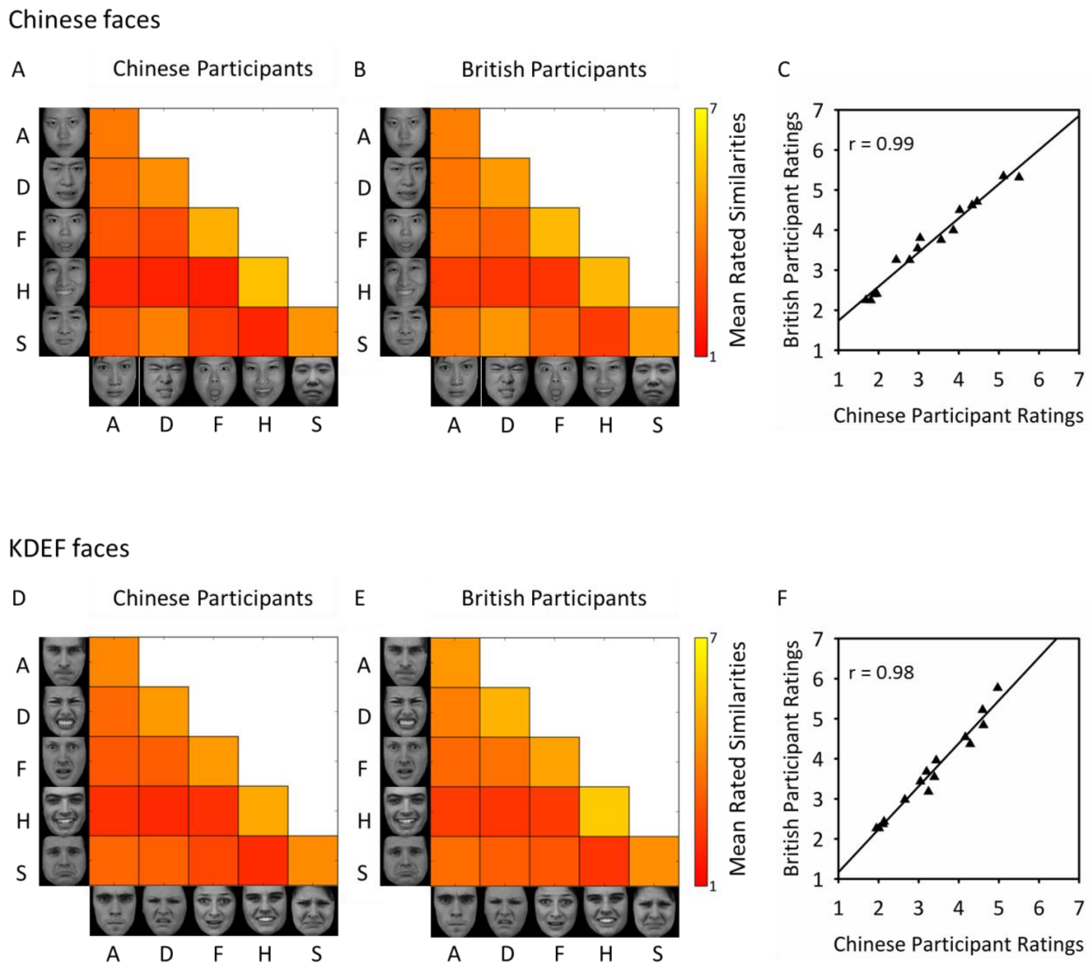


Figure 2.9 Correlation analyses of similarity ratings between Chinese and British participants in Experiment 3.Similarity matrices for (A) Chinese and (B) British participants for presentations of Chinese faces (A: anger, D: disgust, F: fear, H: happy, S: sad). (C) Scatterplot of rating correlation between the two groups for Chinese faces. Similarity matrices for (D) Chinese and (E) British participants for presentations of KDEF faces. (F) Scatterplot of rating correlation between the two groups for KDEF faces. All Chinese face images are from the Chinese Facial Affective Picture System (CFAPS) (Wang & Luo, 2005; Gong et al., 2011) and all Western Caucasian faces are from the Karolinska Directed Emotional Faces (KDEF) (Lundqvist et al., 1998).

As had been noted in Experiment 1, the correlations for the between-category pairs were still very high when the same-expression pairs were removed from the analyses. The correlation was 0.98 ($p < .001$) for the Chinese faces and 0.97 ($p < .001$) for the KDEF faces.

A 5 (Expression: anger, disgust, fear, happiness, sadness) × 2 (Face: Chinese, KDEF) × 2 (Group: Chinese, British) mixed-ANOVA of the arcsin transformed accuracies was conducted to analyse the categorization data. The most important result was a significant two-way interaction of Face × Group, $F(1,38) = 26.11$, $p < .001$, partial $\eta^2 = 0.41$. Further analyses showed that British participants were better at recognizing KDEF (Western appearance) faces than Chinese faces ($t(19) = 5.67$, $p < .001$), whereas Chinese participants showed no difference for categorizing either Chinese or KDEF faces ($t(19) = 1.56$, $p > .1$). However, this interaction could also be decomposed in another way. That is, when comparing the recognition rates between two groups of participants for each face set, our results showed that Chinese participants were slightly better at categorizing Chinese facial expressions than British participants ($t(38) = 1.92$, $p = .06$), while British participants performed slightly better at identifying KDEF facial expressions than Chinese participants ($t(38) = 1.89$, $p = .07$) (see Figure 2.10). This two-way interaction was also qualified by a three-way interaction of Expression × Face × Group, $F(4,152) = 7.02$, $p < .001$, partial η^2 = 0.16. Further analyses revealed that for the Chinese faces, Chinese participants' recognition accuracies were marginally higher than those of British participants for anger ($t(38) = 1.89$, $p = .07$), disgust ($t(38) = 1.77$, $p = .09$), and sadness ($t(38) = 1.95$, $p = .06$) expressions, while for the KDEF faces, British participants were better at recognizing anger expressions than Chinese participants ($t(38) = 3.35$, $p < .01$).

The main effects of Expression ($F(4,152) = 60.56$, $p < .001$, partial $\eta^2 = 0.61$) and Face ($F(1,38) = 8.25$, $p < .01$, partial $\eta^2 = 0.18$), and the interaction of Expression × Face ($F(4,152) = 39.65$, $p < .001$, partial $\eta^2 = 0.51$) were also significant. Further analyses showed that anger ($t(39) = 2.90$, $p < .01$) and sadness ($t(39) = 8.91$, $p < .001$) were better recognized from the KDEF faces than the Chinese faces, whereas fear ($t(39) = 7.22$, $p < .001$) was better identified from the Chinese faces than the KDEF faces. No other significant effects were found.

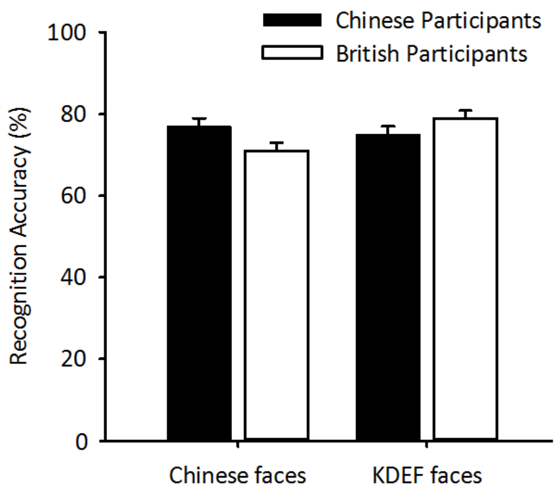


Figure 2.10 Overall recognition accuracies (with standard error bars) for Chinese and British participants from the Chinese and KDEF faces in Experiment 3, plotting the statistically significant Group x Face interaction ($p < .001$).

For the RT analyses, a borderline significant ($p = .06$) main effect of Group was found in the perceptual similarity task, with Chinese participants showing an overall tendency toward slower rating responses than British participants. However, no significant interactions with Group were found, indicating that any overall RT differences were not modified by the own-group or other-group status of the rated faces. In the categorization task, no significant main effect or interactions with Group were found. These results indicated that there were no time differences for participants in processing expressions of their own- and the other-group, and also there was no speed-accuracy trade-off in the categorization task.

The time that the Chinese participants had been in the UK varied from half a month to six years, and was less than a year for 26 out of 40 participants. The same analyses as those used for Experiment 1 were conducted to investigate the potential effect of the time of living in a western environment on participants' processing of facial expressions. In the perceptual similarity task, the correlations between time in the UK and the performance of Chinese participants for the Chinese and KDEF faces were -0.43 (p = .06) and -0.03 (p = .89), respectively. Even though the correlation between time living in the UK of Chinese participants and their performance for Chinese faces was borderline significant, it was the only significant result found in this study, and the relationship between the two was actually in a different direction to that predicted. For the categorization task, the correlations between the time in the UK and the performance of Chinese participants for the Chinese and KDEF faces were 0.20 (p = .40) and 0.33 (p = .15), respectively. These results again indicated that the time spent living in the UK has little effect on Chinese participants' processing of facial expressions.

2.4.3 Discussion

In this experiment, participants performed perceptual similarity and emotion categorization tasks with a full crossover design involving sets of Chinese and western (KDEF) faces that could better represent the range of facial expressions we might encounter in everyday life. Our results confirmed and extended the findings of Experiment 1. In the perceptual similarity task, correlations of performance between the two groups of participants were consistently high for both the Chinese and KDEF faces, indicating no differences in the pattern of perceived similarity between facial expressions across Chinese and British participants. In the categorization task, however, both groups of participants showed marginally higher recognition accuracies for facial expressions expressed by members of their own ethnic group.

The categorization data provide further evidence supporting findings of own-group advantages in recognizing facial expressions of basic emotions. Taken together with the perceptual similarity data, they suggest that the cause of this own-group advantage is to be found in classificatory mechanisms rather than perception per se. Following the logic used for Experiment 2, we then sought to investigate whether this cultural difference in categorization involves differential reliance on information from different parts of the face.

2.5 Experiment 4

Chinese and British participants carried out version (c) of the perceptual similarity task (successive presentation with a mask) and the categorization task with the upper (eyes and eyebrows) or lower (mouth) regions of the Chinese and KDEF stimuli.

2.5.1 Methods

2.5.1.1 Participants

Twenty Chinese (mean age, 21.8 years old) and 20 British (mean age, 19.5 years old) participants were recruited to do the perceptual similarity task and categorization task with only the presentation of the upper half faces, while another 20 Chinese (mean age, 22.65 years old) and 20 British (mean age, 19.35 years old) participants did the same two tasks viewing only the lower half faces. All participants were given a small payment or course credit.

2.5.1.2 Stimuli

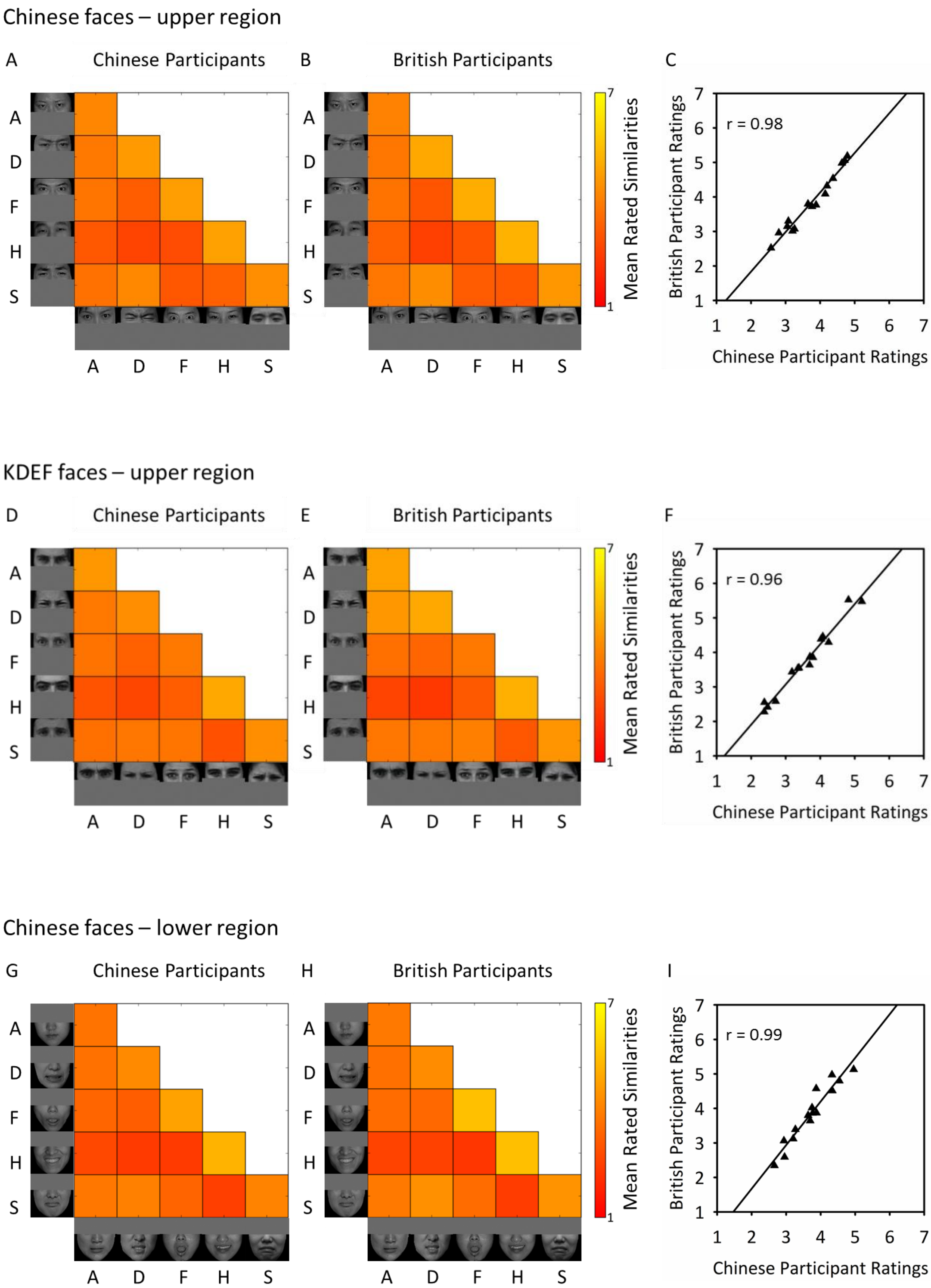
The 25 Chinese and 25 KDEF faces used in the perceptual similarity task in Experiment 3 were used in this experiment for both the perceptual similarity task and the categorization task, except that all face images were divided into upper and lower half regions. The upper region faces were presented with a grey mask covering the lower part, and the lower region faces were presented with a grey mask covering the upper half.

2.5.1.3 Procedure

All participants had to complete the perceptual similarity task first and then the categorization task. In order to minimise the effect of task fatigue, the experiment was divided into two sessions. One half of participants did the two tasks with the Chinese faces first, and then came separately to do the session with the KDEF faces. The other half of the participants did the two tasks with the KDEF faces in the first session and then the Chinese faces in a second session. The 25 face images in each set (5 models posing 5 basic emotions) were presented twice each for each participant in a random order, yielding a total of 50 trials. All other details were identical to those in Experiment 3. As for the other experiments, all Chinese participants showed correct understanding of the meaning of the 5 emotion labels used in the categorization task.

2.5.2 Results

We conducted the same analyses as those used for the other experiments. In the perceptual similarity task, for the upper region, the correlation between rated similarities across all expression pairs between the two groups of participants was very high for the Chinese faces, $r = 0.98$, $p < .001$, and for the KDEF faces, $r = 0.96$, $p < .001$. For the lower region, the correlation between rated similarities across all expression pairs between the two groups of participants was also very high for the Chinese faces, $r = 0.99$, $p < .001$, and for the KDEF faces, $r = 0.98$, $p < .001$. The average ratings for each group of participants and scatter plots are shown in Figure 2.11.



Chinese faces – upper region
A
Chinese Participants
B
British Participants
C
A D F H S
Mean Rated Similarities
7
1
r = 0.98
British Participant Ratings
Chinese Participant Ratings
KDEF faces – upper region
D
Chinese Participants
E
British Participants
F
r = 0.96
Chinese faces – lower region
G
Chinese Participants
H
British Participants
I
r = 0.99

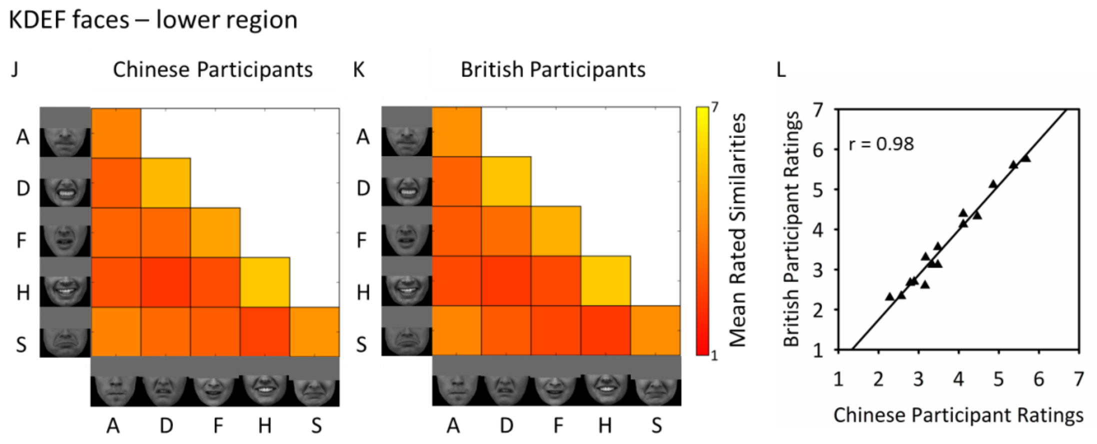


Figure 2.11 Correlation analyses of similarity ratings between Chinese and British participants in Experiment 4. Similarity matrices for (A) Chinese and (B) British participants for presentations of Chinese upper face region (A: anger, D: disgust, F: fear, H: happy, S: sad). (C) Scatterplot of rating correlation between the two groups for Chinese upper region images. Similarity matrices for (D) Chinese and (E) British participants for presentations of KDEF upper face region images. (F) Scatterplot of rating correlation between the two groups for KDEF upper face region. Similarity matrices for (G) Chinese and (H) British participants for presentations of Chinese lower face region images. (I) Scatterplot of rating correlation between the two groups for Chinese lower region images. Similarity matrices for (J) Chinese and (K) British participants for presentations of KDEF lower face region. (L) Scatterplot of rating correlation between the two groups for KDEF lower region images. All Chinese face images are from the Chinese Facial Affective Picture System (CFAPS) (Wang & Luo, 2005; Gong et al., 2011) and all Western Caucasian faces are from the Karolinska Directed Emotional Faces (KDEF) (Lundqvist et al., 1998).

The correlations were still remarkably high for the between-category expression pairs. For the upper region images, the correlation was 0.96 ($p < .001$) for the Chinese faces, and 0.95 ($p < .001$) for the KDEF faces. For the lower region images, the correlation was 0.98 ($p < .001$) for the Chinese faces, and 0.93 ($p < .001$) for the KDEF faces.

In the categorization task, a 5 (Expression: anger, disgust, fear, happiness, sadness) × 2 (Face: Chinese, KDEF) × 2 (Region: upper, lower) × 2 (Group: Chinese, British) mixed-ANOVA of the arcsin transformed accuracy data found a significant interaction of Face × Region × Group ($F(1,76) = 8.53$, $p < .01$, partial $\eta^2 = 0.10$). This three-way interaction forms the principal result of interest. Decomposition of the interaction showed that for the upper half images, there was no difference between the two participant groups either for the Chinese faces ($t(38) = 0.60$, $p > .1$) or for the KDEF faces ($t(38) = 0.58$, $p > .1$). However, for the lower half images, Chinese participants showed better overall performance than British participants at identifying the Chinese images ($t(38) = 2.41$, $p < .001$), whereas British participants were better overall at identifying the lower parts of the KDEF images than Chinese participants ($t(38) = 2.93$, $p < .001$) (see Figure 2.12).

We also found significant main effects of Expression ($F(4,304) = 118.53$, $p < .001$, partial $\eta^2 = 0.61$), Face ($F(1,76) = 64.15$, $p < .001$, partial $\eta^2 = 0.46$), and Region ($F(1,76) = 32.16$, $p < .001$, partial $\eta^2 = 0.30$). In addition to these main effects, the analysis also revealed significant two-way interactions of Expression × Region ($F(4,304) = 8.81$, $p < .001$, partial $\eta^2 = 0.10$), Expression × Group ($F(4,304) = 2.92$, $p < .05$, partial $\eta^2 = 0.04$), Expression × Face ($F(4,304) = 34.15$, $p < .001$, partial $\eta^2 = 0.31$), Face × Group ($F(1,76) = 19.37$, $p < .001$, partial $\eta^2 = 0.20$), and Face × Region ($F(1,76) = 12.53$, $p < .001$, partial $\eta^2 = 0.14$). Four significant three-way interactions were also found: (1) Expression × Region × Group, $F(4,304) = 2.55$, $p < .05$, partial $\eta^2 = 0.03$. Further analyses showed that for the upper region faces, Chinese participants were better at identifying happy expressions than British participants ($t(38) = -3.30$, $p < 0.01$), while for the lower region faces, British participants were better at identifying disgust expressions than Chinese participants ($t(38) = 2.71$, $p < .05$). (2) Expression × Face × Group, $F(4,304) = 4.66$, $p < .001$, partial $\eta^2 = 0.06$. Further analyses showed that Chinese participants were better at recognizing Chinese anger expressions than British participants ($t(78) = 3.69$, $p < .001$). (3) Expression × Face × Region, $F(4,304) = 3.27$, $p < .01$, partial $\eta^2 = 0.04$. Further analyses showed that for the upper region faces, anger ($t(39) = 5.30$, $p < .001$) and sadness ($t(39) = 6.37$, $p < .001$) expressions were better detected from the KDEF faces than the Chinese faces, while fear expressions were better identified from the Chinese faces than the KDEF faces ($t(39) = 4.70$, $p < .001$). For the lower region faces, anger ($t(39) = 5.06$, $p < .001$), disgust ($t(39) = 5.16$, $p < .001$) and sadness ($t(39) = 5.92$, $p < .001$) expressions were all found to be better recognized from the KDEF faces than the Chinese faces. The four-way interaction between Expression × Face × Region× Group was not significant, $F(4,304) = 0.85$, $p > .1$, partial $\eta^2 = 0.01$. The rest of the effects were not significant, either.

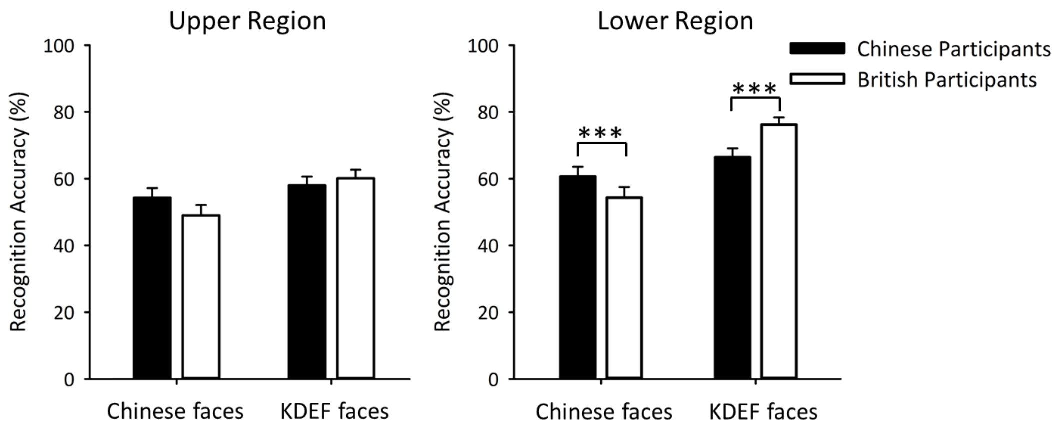


Figure 2.12 Overall emotion recognition accuracies (with standard error bars) for Chinese and British participants from upper and lower regions of the Chinese and the KDEF faces in Experiment 4. Asterisks denote conditions with a significantly higher recognition rate in comparison with the corresponding paired condition from the lower face region. ***: $p < .001$.

For the RT analyses, no significant main effects or interactions with Group were detected for either the perceptual similarity task or the categorization task. These results again indicated that there were no time differences for participants in processing expressions of the own- and the other-group, and also there was no speed-accuracy trade-off in the categorization task.

The time that the Chinese participants had been living in the UK varied from half a month to six years, and was less than a year for 35 out of 40 participants. Correlations of the time living in the UK for Chinese participants and their performance showed no significant effects (Upper region perception: for Chinese faces, r = 0.01, p = .98; for KDEF faces, r = -0.02, p = .92. Upper region categorization: for Chinese faces: r = -0.09, p = .71; for KDEF faces: r = -0.27, p = .26. Lower region perception: for Chinese faces, r = 0.14, p = .55; for KDEF faces, r = 0.11, p = .65. Lower region categorization: for Chinese faces: r = 0.12, p = .62; for KDEF faces: r = 0.02, p = .95).

2.5.3 Discussion

In this study, we further extended the results from Experiment 2 through the use of upper and lower parts of Chinese and Western (KDEF) faces. In the perceptual similarity task, correlations between the ratings of Chinese and British participants always showed high consistencies even with the presentation of half faces (either upper or lower), confirming that there was no cultural difference in patterns of perceptual similarity between facial expressions of basic emotions across Chinese and British participants. However, in the categorization task, an own-group advantage was detected with the presentation of the lower region of the faces. Both groups of participants were better at recognizing facial expressions expressed in the mouth region by members of their own cultural group. In contrast, no significant differences in overall categorization accuracy were found between the two cultural groups with the upper face region. These results clarify the own-group advantage found in Experiment 3, by demonstrating that cultural differences in categorization of facial expressions are mainly linked to differences in information decoding in the lower region of faces. Moreover, they replicate the finding from Experiment 2 that cultural differences in categorization accuracy largely involve the lower (mouth) region of the face and extend this by showing a crossover own-group advantage in which Chinese participants can make better use of information from the mouth region of Chinese than Western faces and British participants can make better use of information from the mouth region of Western than Chinese faces.

2.6 General Discussion

In the four experiments reported here, we systematically examined cultural similarities and differences in the perception and categorization of facial expressions of basic emotions between Chinese and British participants. To the best of our knowledge, our study is the first to systematically examine cultural similarities and differences in both the perception and categorization of facial expressions of basic emotions between Chinese and British participants. Our results revealed a clear difference between the influences of culture on the way in which people perceive and categorize facial expressions. In our perceptual task, participants rated the similarity between facial expressions of basic emotions posed by two different individuals, so that differences in identity had to be ignored to make the perceptual judgement. In terms of perceiving facial expressions, we found no group differences in the patterns of inter-expression similarity; correlations between Chinese and British participants for the rated perceptual similarities between pairs of expressions were always high across four experiments. In terms of categorizing expressions, however, participants showed a small but statistically reliable advantage for facial expressions expressed by members of their own cultural group than those expressed by others. These categorization results replicate those of previous studies showing that there is an own-group advantage in recognizing facial expressions (Izard, 1971; Ekman, 1972; Ekman et al., 1969; Jack et al., 2009). The results from the perceptual task constrain the possible interpretations of this own-group categorization advantage.

In addition, we further investigated whether there are cultural differences in processing strategies or biases involving different parts of the face between Chinese and British participants. This was based on results of previous studies suggesting that people from East Asian and Western cultures tend to focus on different facial signals in recognizing and even internally representing facial expressions (Jack et al., 2009; Jack et al., 2012). To address this question, we repeated the perceptual similarity and the categorization tasks, but with the presentation of only the upper (eyes, eyebrows and forehead) or lower (mouth and chin) part of each face. We still did not find any group differences in the patterns of similarity ratings for pairs of expressions between Chinese and British participants, for either upper or lower parts of the face. These data are therefore in line with our conclusion that there is no group difference in perception of facial expressions, and demonstrate that this lack of a basic perceptual difference extends to the perception of local features (such as eyes or mouth).

The results from the categorization task with part faces offered an interesting contrast. The own-group advantage in recognizing facial expressions between Chinese and British participants only reached statistical significance with the presentation of the lower region of each face; no significant own-group advantage was found for the upper region that includes the eyes and eyebrows. These results differ from Jack et al.'s (2009, 2012) view that East Asian participants do not make much use of the mouth region in recognizing facial expressions. Instead, we found that participants with either Chinese or Western cultural backgrounds could make use of information from the mouth region, but both groups were slightly better at using it to recognize facial expressions posed by members of their own ethnic group.

Even though our main focus was on investigating the perceived similarity ratings of expressions for participants and the mean accuracies for identifying expressions, we also ran analyses on the response times to see whether there were differences in the time required for participants to process faces of their own- and the other-group. Our results, however, indicated that participants from the Chinese and British cultures spent the same amount of time in processing faces of either their own group or the other group. There was no evidence of speed-accuracy tradeoffs, or a general tendency to spend longer in evaluating own-group faces.

Neuroimaging studies have indicated that the posterior superior temporal sulcus (pSTS) and the amygdala respond to different types of change in facial expressions. The amygdala is more sensitive to the categorical representation of facial expressions, whereas the pSTS uses a more continuous representation (Harris et al., 2012; 2014a). These findings suggest that it may be possible to further investigate the dissociation between perception and categorization of facial expressions between Chinese and British participants at the neural level.

A particular strength of the present study was that we were able to include a fully balanced design in Experiments 3 and 4, with Chinese and Western (KDEF) faces viewed by both Chinese and British participants. We also took care to use both tightly standardized (Experiments 1 and 2, with Ekman and Friesen faces based on muscle action units), and also more naturally variable sets of images (the Chinese and KDEF faces used in Experiments 3 and 4 were both made by asking actors to imagine emotional scenarios).

However, as part of this design we used the English labels for the five basic emotions for all participants in the categorization task. This meant that our Chinese participants were not performing the categorization task in their native language, and we therefore included an additional task to confirm their correct understanding of the English emotion words. Our reason for not translating the basic emotion labels into Chinese was that studies have shown that some cultural differences in emotion recognition might be attributable to differences in the way that the vocabularies of some languages are tailored to conceptualising some emotions (Matsumoto & Assar, 1992). Such differences could introduce confounds into the design if we had translated the labels into Chinese, and we therefore preferred to keep the task consistent across the two groups of participants by using the English labels. We think this decision was justified on the basis that we detected no group differences in RTs between Chinese and British participants in the categorization tasks. Moreover, the use of English labels would in any case only be a potential problem for Chinese participants, and therefore cannot explain the observed interactions in emotion categorization accuracy between the participant group and the own-group or other-group status of the stimulus face.

How then is the own-group advantage in categorizing expressions to be explained? Two points stand out from our data. First, though reliable, the advantage is not large, and it does not sit easily with the idea of substantial inter-cultural differences in categorization style (Nisbett & Masuda, 2003). Instead, consistent with the idea of universality (Darwin, 1872; Ekman, 1980), there is no sense in which our participants were 'blind' to the expressions of someone from another culture. Second, we found no evidence that the own-group advantage reflects any more fundamental perceptual difference.

A number of ideas have been offered in the literature to try to explain the own-group advantage in recognizing facial expressions. For instance, it might be caused by cultural differences in display and decoding rules regarding facial expressions (Ekman & Friesen, 1969; Matsumoto & Ekman, 1989), or from variations in the way of encoding across cultures (Elfenbein & Ambady, 2002b, 2003c). Such explanations imply that observers should more effectively understand emotions expressed by members of a cultural group to which they have had significant exposure. Elfenbein and Ambady (2003c) found that Chinese students who had been living in the USA for an average of 2.4 years could better at recognizing facial expressions of members of their host culture than those of their own-group members, indicating that cultural familiarity could occur within this overall time period. In the present study, we also examined effect of time spent in the UK by our Chinese participants on their perception and recognition of facial expressions. However, we did not find any significant correlations between the time of staying in the UK and participants' performance in the two tasks. Two reasons might explain the discrepancy between our results and those of Elfenbein and Ambady (2003c): (1) Even though the time our Chinese participants had been in the UK varied from one month to almost nine years, many of them had been living in the UK for less than a year. (2) As Elfenbein and Ambady (2003c) argued, the own-group advantage in emotion recognition accuracy may vary according to the level of exposure to the other-group culture, which is difficult to measure.

Our findings extend our understanding of the similarities and differences in the way people from different cultures perceive and recognize facial expressions, and constrain the possible interpretations of the own-group advantage in facial expression recognition. A highly relevant theoretical debate has arisen from studies of the own-group advantage found in many previous studies of face identity recognition (Bothwell, Brigham, & Malpass, 1989; Shapiro & Penrod, 1986). Accounts of this own-group advantage in identity recognition have either emphasized perceptual learning, because the cues that best serve to identify individuals may differ between faces of different ethnicities, or emphasized social psychological processes because participants may be less motivated to individuate faces they see as belonging to an 'out-group' (Meissner & Brigham, 2001; Pettigrew & Tropp, 2000; Hugenberg, Young, Bernstein, & Sacco, 2010). Our finding that own-group advantages in facial expression categorization were largely restricted to the lower part of the face makes the social psychological type of explanation an unlikely candidate here; both the upper or lower parts of the faces are of 'Western' or 'Chinese' appearance, but only the lower part leads to a categorization advantage. It therefore seems more likely that our findings reflect relatively minor cultural 'stylistic' differences in the way in which these emotions are expressed around a common overall template, and we note of course that the organization of the facial muscles makes the lower part of the face relatively mobile compared to the more limited range of movements possible in the eye region, and hence more capable of developing such differences. Above all, though, the fact that the own-group categorization advantage is small in comparison to the level of cross-cultural agreement implies that the idea of universality should not be hastily rejected.

3. Differences in holistic processing do not explain cultural differences in the recognition of facial expression (accepted by the Quarterly Journal of Experimental Psychology)

3.1 Introduction

The question of whether a small number of facial expressions correspond to basic emotions with a long evolutionary history, and hence are universally recognised, has elicited considerable debate since Darwin (1872) put forward the suggestion in the nineteenth century. From the research stimulated by this debate, two consistent findings stand out. First, recognition of facial expressions of basic emotions is substantially above-chance in all cultures tested to date (Biehl, et al., 1997; Ekman, 1972; Izard, 1971); this finding is consistent with the universality hypothesis. Second, although always above-chance, there are none the less some cultural differences and people are often better at recognizing expressions posed by their own-race versus other-race members (Elfenbein & Ambady, 2002b; Jack et al., 2012; Yan, Andrews, & Young, 2016a; Yan, Andrews, Jenkins, & Young, 2016b); these findings of cross-cultural differences and own-group advantages set limits on the extent of universality.

A key unresolved issue concerns what causes cultural difference in facial expression recognition. A novel hypothesis proposed by Jack and colleagues (2012) suggests that the differences between cultural groups are driven by people from different cultural backgrounds paying attention to different facial signals when processing facial expressions. For example, in a study that used reverse correlation methods to estimate the internal representation of static facial expressions, Jack et al. (2012) maintained that East Asian participants mainly use information from the eye region when processing facial expressions, whereas Western Caucasian participants rely more evenly on both the eye and mouth regions. From this perspective, the cross-cultural differences reflect underlying differences in mental representations resulting from differences in the attended regions of the face. A recent study by Yan et al. (2016a) therefore systematically investigated cross-cultural similarities and differences in the perception as well as the recognition of facial expressions of five basic emotions (anger, fear, happiness, disgust, and sadness). By asking Western Caucasian and Chinese participants to make similarity ratings to pairs of expressions or to identify the emotion from facial expressions, Yan et al. (2016a) showed that there was actually considerable consistency in the way each group of participants perceived facial expressions, but a small cross-cultural difference in recognizing facial expressions which was driven in part by an own-group advantage in recognizing anger and disgust.

Although their findings offered at best limited support for Jack et al.'s (2012) claim of an underlying difference in perceptual representations, one limitation of Yan et al.'s (2016a) study was that the most confusable expressions they used were anger and disgust, so that it was unclear whether the own-group advantage Yan et al. (2016a) found for recognizing anger and disgust reflected something to do with expressions of these emotions *per se*, or simply the fact that they were the most confusable expressions in the set investigated (happiness, sadness, fear, anger, disgust). In the present Experiment 1, we therefore added facial expressions of surprise to the set used by Yan et al. (2016a). In studies of facial expression recognition, surprise is confused with fear more often than anger is confused with disgust (Calvo & Lundqvist, 2008; Ekman & Friesen, 1976; Palermo & Coltheart, 2004; Wiggers, 1982). Hence including expressions of surprise as well as fear allows us to test whether the own-group advantage is driven by overall confusability (in effect, by task difficulty). Moreover, facial expressions of surprise were also included in Experiment 1 because Jack et al. (2012) have argued that surprise plays an important role in driving the group differences in expression perception.

Jack et al.'s (2012) hypothesis that East Asian participants mainly use information from the eye region to recognize facial expressions also predicts that holistic processing of expression should be reduced in comparison to Western Caucasian participants. For Western participants it is well-established that facial expressions are perceived holistically, with information from the mouth region modifying the interpretation of information from the eye region and vice versa. The most well-known demonstration involves a facial expression variant of the face composite paradigm devised by Young, Hellawell and Hay (1987). Calder and colleagues (2000) created images that combined the upper half of one expression with the lower half of a different expression. They found that participants were slower at identifying expressions from either the upper or the lower part of these images when the two half parts were presented in a face-like aligned composite format than when the same parts were presented in a misaligned format that was not face-like. This effect has been replicated in other studies of Western participants (Flack et al., 2015; Tanaka, Kaiser, Butler, & Le Grand, 2012). It is interpreted as indicating that holistic perception of the face-like aligned composite stimuli makes it difficult for participants to ignore information from the irrelevant part of the image (i.e. to ignore information from the bottom half when classifying the top half, or vice versa), In contrast, because the misaligned stimuli do not create a face-like configuration, they are not susceptible to this holistic interference.

In Experiment 2 we therefore tested the expression composite effect in Western Caucasian and East Asian participants, using a paradigm modelled on Calder et al. (2000). If the recognition of expressions by East Asian participants is dominated by information from the eye region, we expect either a reduced composite effect overall or a reduced effect when it is the part of the face containing the eye region that has to be classified. An additional reason for testing the expression composite effect cross-culturally is that some studies have linked own-group advantages in the recognition of facial identity (rather than expression) to holistic processing (Tanaka, Kiefer, & Bukach, 2004; Michel, Caldara, Rossion, 2006a; Michel, Rossion, Han, Chung, & Caldara, 2006b). However, findings of enhanced holistic processing of own-race faces are by no means consistently obtained (Hayward, Crookes, & Rhodes, 2013) and no studies have yet looked at cross-cultural differences in holistic processing of facial expressions.

3.2 Experiment 1

This experiment examined cross-cultural similarities and differences in perceiving and recognizing facial expressions of six basic emotions with a full crossover design that included Chinese and Western faces and Chinese and Western participants. Separate perceptual similarity and emotion categorization tasks were used, with the perceptual task asking participants to rate the similarity of facial expressions across pairs of face photographs and the categorization task involving forced-choice recognition of the facial expressions. This experiment also aimed to investigate whether the own-group advantages in expression recognition found by Yan and colleagues (2016a) was driven by certain confusable emotion categories. Studies have found that there are confusions among certain emotion categories, such as anger and disgust, and fear and surprise (Calvo & Lundqvist, 2008; Ekman & Friesen, 1976; Palermo & Coltheart, 2004; Wiggers, 1982). We were interested in whether cultural differences in expression recognition might largely be driven by confusions between these emotions. In addition, the inclusion of facial expressions of surprise is of interest because, according to Jack et al., (2012), there are particularly clear cultural differences in the mental representation of surprise.

3.2.1 Methods

3.2.1.1 Participants

Eighteen Chinese students brought up in China with Chinese parents (13 females; mean age, 21.4 years) and 18 Western Caucasian students brought up in western countries with Western Caucasian parents (14 females; mean age, 20.8 years) were recruited from the University of York. All participants gave their written consent prior to the experiment. The University of York Department of Psychology Ethics Committee approved the study.

3.2.1.2 Stimuli

Photographs of facial expressions of six basic emotions (anger, disgust, fear, happiness, sadness, and surprise) were selected from two face sets; the Chinese Facial Affective

Picture System (CFAPS) (Wang & Luo, 2005; Gong, Huang, Wang, & Luo, 2011) posed by Chinese models, and the Karolinska Directed Emotional Faces (KDEF) (Lundqvist, Flykt, & Öhman, 1998) posed by Western Caucasian models. In total, 120 Chinese and 120 Western Caucasian faces (with 20 exemplars of each of the 6 emotions) were used for the categorization task, and 18 Chinese and 18 Western Caucasian images (3 exemplars of each of the 6 emotions) were used for the perceptual similarity task.

All images were converted to greyscale and cropped to remove hairstyles and background as far as possible. When viewed in the experiment each image subtended a visual angle of approximately 7 x 8 degrees. Figure 3.1 shows examples of images used in the experiment. The images for five of the basic emotions (anger, disgust, fear, happiness, and sadness) were the same as those previously used by Yan et al. (2016a).

Figure 3.1 Example face images for 6 emotions posed by different models from the Chinese Facial Affective Picture System (CFAPS; Wang & Luo, 2005; Gong et al., 2011) and the Karolinska Directed Emotional Faces (KDEF; Lundqvist et al., 1998).

3.2.1.3 Procedure

Participants viewed expression images using a computerized task programmed with PsychoPy software (www.psychopy.org). All participants completed the perceptual similarity rating task first, and then the forced-choice expression categorization task.

In the perceptual similarity task, participants saw two facial expressions posed by different actors presented simultaneously side by side for 1.5 seconds. Their task was to rate the similarity of the expression pairs on a 7-point scale, with 1 indicating not very similar expressions and 7 very similar expressions. There were 15 different types of expression pairings in which a photograph showing one expression was always paired with a photograph showing a different expression (e.g. anger with disgust, anger with fear, anger with surprise, and so on; resulting in 15 possible types of combination). Same expression pairs (e.g. anger with anger, disgust with disgust) were not included because Yan et al. (2016a) found that these always generated high rated similarities. We therefore chose to focus on the perceived similarity of between-expression pairs, which offer a stronger test of whether differences between expressions are perceived equivalently across cultures. Because each emotion expression was posed by 3 actors, there were a total of 9 possible combinations for each of the 15 expression pairs, leading to a total of 135 trials for each set of faces. Ten additional practice trials were included to familiarize the participants with the task prior to the formal experiment. The trial order was random across participants.

In the categorization task, participants only saw one face at each time and they had to perform a six-alternative forced-choice task (6AFC) to identify its facial expression as happiness, sadness, anger, disgust, fear, or surprise. Each face was presented for 1 second, and the participants were asked to make their response as quickly and as accurately as possible. Responses were made via keypresses 1-6 for the expressions and the mapping between emotion labels and keys was counterbalanced across participants. The code for keypresses was always visible on screen. There were a total of 120 trials with Chinese faces and 120 trials with Western Caucasian faces, with each set split randomly into two blocks. Participants saw the face images in a block order of either 'Chinese-Caucasian-Chinese-Caucasian' or 'Caucasian-Chinese-Caucasian-Chinese', which was counterbalanced across participants. There was also a 10-trial practice session at the beginning.

After these two tasks, all Chinese participants were asked to write down the Chinese names of the six emotion labels used in the categorization task, to check comprehension of the English words. Two native Chinese speakers verified that the labels were all correctly understood by the Chinese participants. They were also asked to fill in a short questionnaire reporting how long they had been in the UK (see Appendix 2, for details).

3.2.2 Results

The experiment involved perceptual similarity rating and forced-choice categorization tasks. We will consider each in turn, looking separately at both accuracies and patterns of confusions in the categorization task.

3.2.2.1 Perceptual Similarity Task

To analyse the similarity ratings for the perceptual similarity task, we followed Yan et al.'s (2016a) procedure of calculating the average similarity ratings for each pair of emotions for each participant (i.e. the average rated similarity of anger-disgust pairs, anger-fear pairs, etc.). The resulting 15 averaged ratings across participants were then used to create perceptual similarity matrices for both the Western Caucasian faces and the Chinese faces in each group of participants. By correlating the values in these similarity matrices across the different participant cultures we can then measure the amount of cross-cultural agreement.

Figure 3.2 shows the similarity rating matrices for Western Caucasian and Chinese faces and Western Caucasian and Chinese participants. The correlations between the similarity rating matrices between Chinese and Western Caucasian participants for both Western Caucasian faces (r = 0.98, p < .001) and for Chinese faces (r = 0.97, p < .001), indicating that the perception of the expressions was highly consistent between Western Caucasian and Chinese participants. These results were consistent with the results found with only 5 emotions by Yan et al. (2016a).

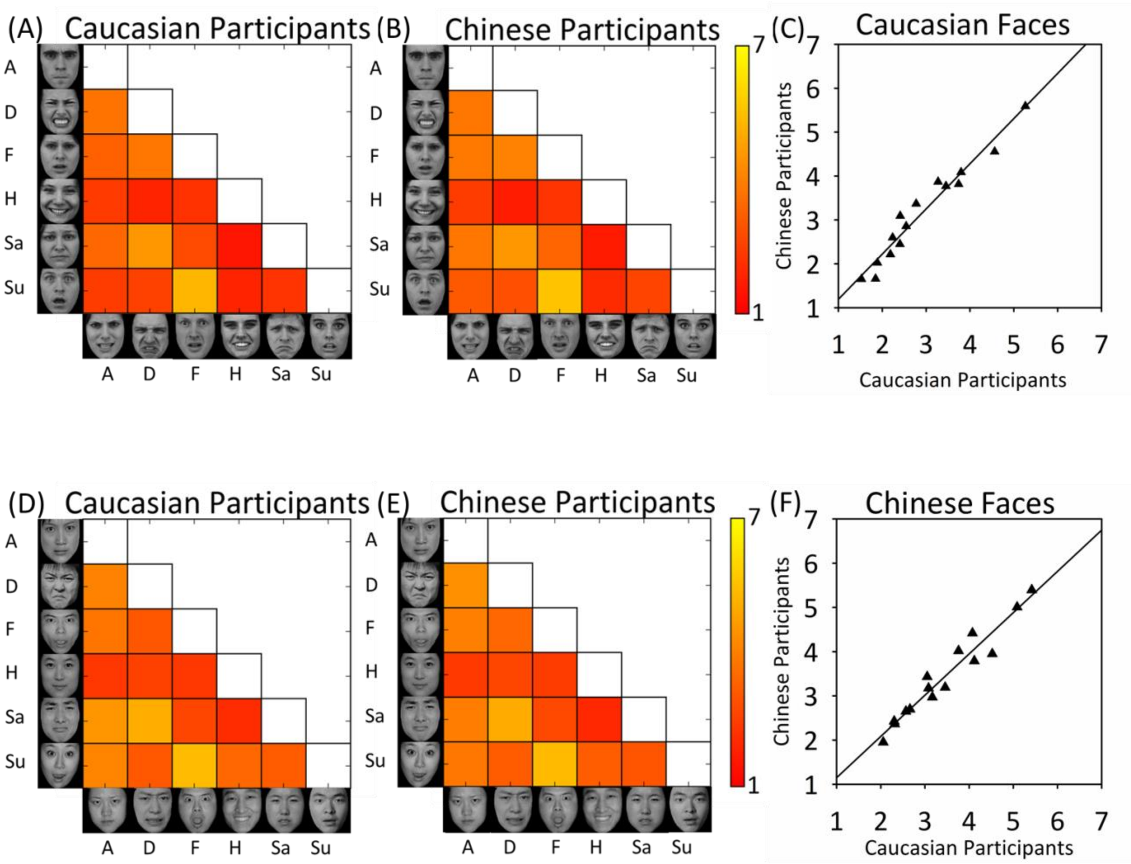


Figure 3.2 Correlation analyses of similarity rating patterns between Western Caucasian and Chinese participants. Perceptual similarity matrices for (A) Western Caucasian and (B) Chinese participants with Western Caucasian faces (A: anger, D: disgust, F: fear, H: happiness, Sa: sadness, Su: surprise). (C) Scatterplot of correlation between two groups of participants with Western Caucasian faces ($r = 0.98$, $p < .001$). Perceptual similarity matrices for (D) Western Caucasian and (E) Chinese participants with Chinese faces. (F) Scatterplot of correlation between two groups of participants with Chinese faces ($r = 0.97$, $p < .001$).

3.2.2.2 Categorization Task

Western Caucasian participants were more accurate in judging facial expressions from Western Caucasian faces (77% ± 1%) compared to Chinese faces (69% ± 1%). In contrast, there was no difference in overall accuracy for Chinese participants judging Western Caucasian (72% ± 1%) or Chinese (72% ± 2%) faces. In addition, the recognition rates for these expressions of Western Caucasian and Chinese faces in two groups of participants were very consistent with the accuracies found in Experiment 3 of Chapter 2 (in which two groups of participants were required to identify facial expressions of five emotions), except fear and surprise (Appendix 3). The recognition rates for fear expressions in Experiment 3 of Chapter 2 were overall higher than the accuracies in this experiment. This is reasonable because surprise was not tested in Experiment 3, which would avoid confusions of surprise with fear (Calvo & Lundqvist, 2008; Ekman & Friesen, 1976; Palermo & Coltheart, 2004; Wiggers, 1982).

A mixed ANOVA was conducted on the arcsine transformed percentage recognition accuracies with Group (Western Caucasian participants, Chinese participants) as a between-subject factor, and Face Ethnicity (Western Caucasian faces, Chinese faces) and Emotion (anger, disgust, fear, happiness, sadness, surprise) as within-subject factors. This showed an own-group advantage in the form of a significant interaction of Face Ethnicity x Group, $F(1,34) = 20.8$, $p < .001$, partial $\eta^2 = 0.38$, shown in Figure 3.3A. Further analyses to decompose this interaction revealed that for the Western Caucasian participants, there were significant recognition accuracy differences between Western Caucasian and Chinese faces, $F(1,34) = 40.1$, $p < .001$, while for the Chinese participants the differences between the two sets of faces were nonsignificant, $F(1,34) = 0.02$, $p > .1$. This interaction was also moderated by a three-way interaction of Emotion x Face Ethnicity x Group, $F(5,170) = 6.1$, $p < .001$, partial $\eta^2 = 0.15$.

To further investigate the potential group differences in each emotion category, we decomposed the three-way interaction to look for a Face Ethnicity x Group interaction separately for each emotion (Figure 3.3B and 3.3C). Our analyses found that the interaction of Face Ethnicity x Group was only significant for anger ($F(1,34) = 36.1$, $p < .001$, partial $\eta^2 = 0.51$) and disgust ($F(1,34) = 5.2$, $p < .05$, partial $\eta^2 = 0.13$). In these significant two-way interactions, there were significant differences between Western Caucasian and Chinese anger faces for both the Western Caucasian participants (who were better at recognizing Western Caucasian expressions, $F(1,34) = 28.5$, $p < .001$) and the Chinese participants (who were better at recognizing Chinese expressions, $F(1,34) = 9.7$, $p < .01$), while the differences between Western Caucasian and Chinese disgust faces only reached significance for Chinese participants ($F(1,34) = 8.6$, $p < .01$).

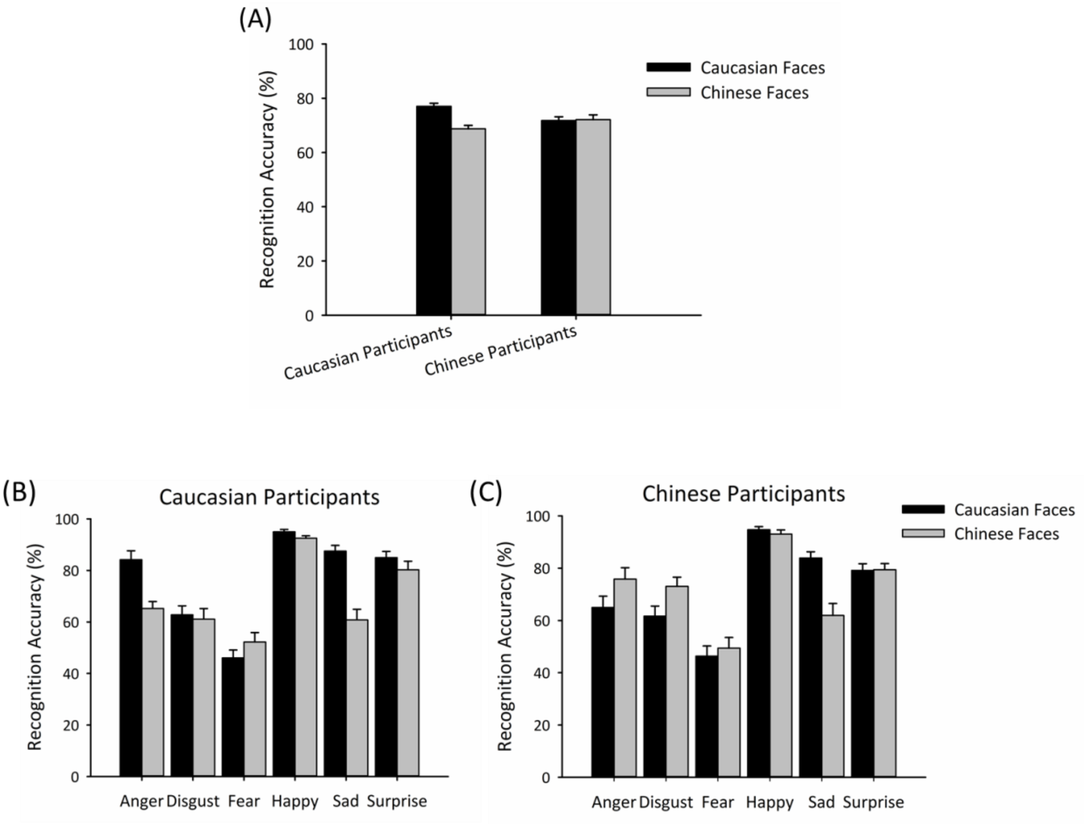


Figure 3.3 Percentage recognition accuracies between Western Caucasian and Chinese participants in Experiment 1. (A) Overall percentage recognition accuracies (with standard error bars) for Western Caucasian and Chinese participants from the Western Caucasian and Chinese facial expressions in the categorization task. (B) (C) Percentage recognition accuracies (with standard error bars) for the six basic emotions by Western Caucasian and Chinese participants presented with Western Caucasian and Chinese facial expressions.

Besides the above results that reflect our main focus of interest, the ANOVA also found significant main effects of Face Ethnicity, $F(1,34) = 19.3$, $p < .001$, partial $\eta^2 = 0.36$, and Emotion, $F(5,170) = 74.4$, $p < .001$, partial $\eta^2 = 0.69$. These main effects were qualified by the interaction of Face Ethnicity x Emotion, $F(5,170) = 17.3$, $p < .001$, partial $\eta^2 = 0.34$, with the Western Caucasian sadness expressions being easier to recognize than Chinese sadness expressions in the sets used, $t(35) = 8.6$, $p < .001$.

We also conducted an equivalent mixed ANOVA on the median reaction times (RTs) for the correct responses in the categorization task. This did not find significant interactions of Face Ethnicity and Group (Face Ethnicity x Group: $F(1,34) = 1.3$, $p > .1$, or Face Ethnicity x Emotion x Group: $F(5,170) = 1.5$, $p > .1$), indicating that there were no cultural differences in response time to facial expressions posed by own- and other-race members, and that there was no speed-accuracy trade-off in the categorization task.

As well as examining categorization accuracies, we also looked at the confusions made by the two groups of participants when identifying facial expressions of the six basic emotions in the categorization task. To do this we created separate confusion matrices for each set of faces (Western Caucasian or Chinese) for each group of participants. These are shown in Figure 3.4. Each matrix represents the pattern of participants' responses, with the y-axis indicating the intended emotion categories and the x-axis indicating participants' responses as the intended or different emotions. In order to compare participants' confusion matrices in the categorization task with their similarity rating matrices from the perceptual similarity task, we averaged the two cells of the same expression pairs (e.g. anger mistaken for disgust and disgust mistaken for anger) in each confusion matrix to create a generic confusion matrix and we also removed the accuracies for intended expressions that fall along the diagonal (i.e. the accuracies for recognizing fear as fear, disgust as disgust and so on). In this way we arrived representations of categorization confusions (Figure 3.4) that were similar in structure to the way we represented the perceptual similarity data (Figure 3.2).

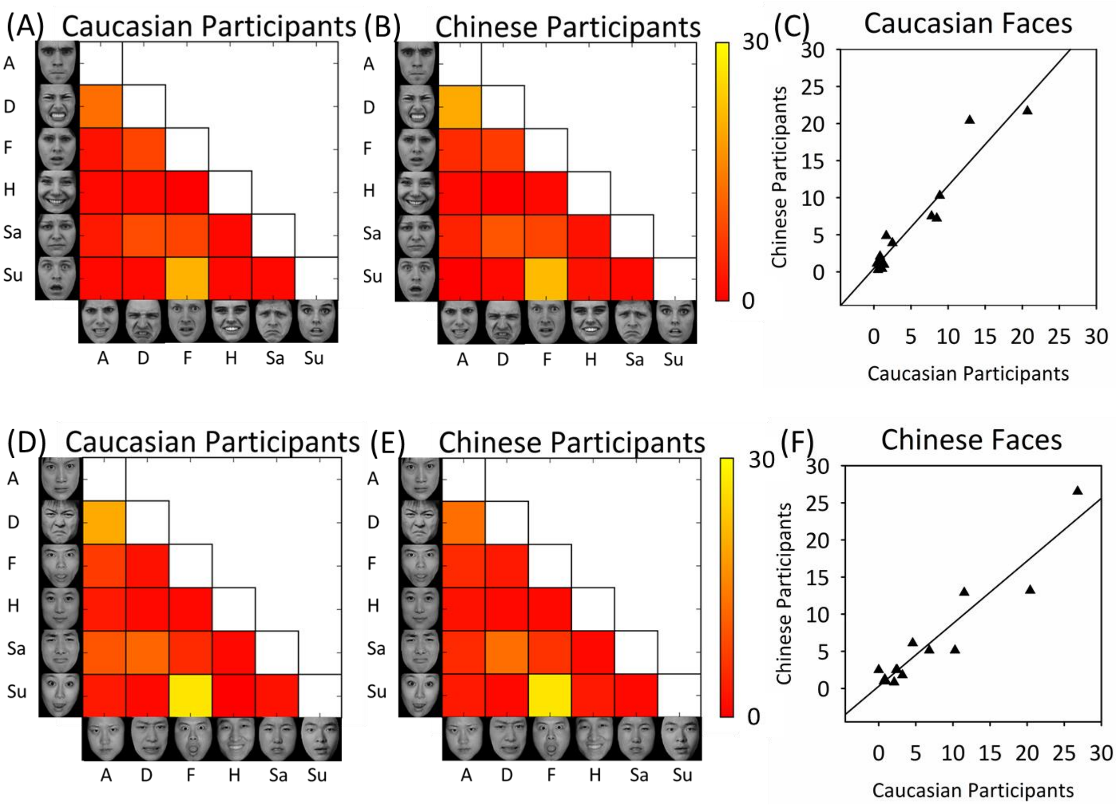


Figure 3.4 Confusion matrices analyses for Western Caucasian and Chinese participants in Experiment 1. Confusion matrices for (A) Western Caucasian and (B) Chinese participants categorizing Western Caucasian faces (A: anger, D: disgust, F: fear, H: happiness, Sa: sadness, Su: surprise). (C) Scatterplot of correlation of the confusion patterns between the two groups of participants with Western Caucasian faces (r = 0.96, p < .001). Both the x- and y-axis indicate the percentage confusion rates of different pairs of expressions. Confusion matrices for (D) Western Caucasian and (E) Chinese participants categorizing Chinese faces. (F) Scatterplot of correlation of the confusion patterns between the two groups of participants with Chinese faces (r = 0.95, p < .001).

We were then able to measure the similarity between these different confusion matrices using correlations, in the same way as we had measured the similarity between the perceptual ratings matrices. Again, the correlation between Chinese and Western Caucasian participants for each set of faces were very high; for Western Caucasian faces, r = 0.96, p < .001, and for Chinese faces, r = 0.95, p < .001, indicating that the overall patterns of confusions between expressions for both Western Caucasian and Chinese participants were very consistent.

As a further step, we also compared the correspondence between the patterns of perceptual similarity ratings shown in Figure 3.2 and the categorization confusion matrices shown in Figure 3.4. Once again we found substantial consistencies between patterns across these two different tasks, indicating that the higher the similarity perceived by the participants for each pair of expressions, the more there were recognition confusions among those expression pairs. The correlations of response patterns between two tasks were: Western Caucasian faces for Western Caucasian participants, r = 0.85, p < .001, Western Caucasian faces for Chinese participants, r = 0.78, p < .001, Chinese faces for Western Caucasian participants, r = 0.76, p < .001, and Chinese faces for Chinese participants, r = 0.82, p < .001.

A noticeable feature of Figure 4 is that the main confusions between expressions involve fear with surprise and anger with disgust. In other studies of facial expression recognition, surprise is confused with fear more often than anger is confused with disgust (Calvo & Lundqvist, 2008; Ekman & Friesen, 1976; Palermo & Coltheart, 2004; Wiggers, 1982). To confirm that this was the case in the present data and explore the impact of face and participant ethnicity, we conducted a further ANOVA with Group (Western Caucasian participants, Chinese participants) as a between-subject factor, and Face Ethnicity (Western Caucasian faces, Chinese faces) and Emotion (the 15 possible emotion pairings, e.g., anger-disgust, anger-fear) as within-subject factors. This revealed a significant three-way interaction of Group x Face Ethnicity x Emotion, $F(14,476) = 6.1$, $p < .001$, partial $\eta 2 = 0.15$. In this interaction the confusions between fear and surprise and anger and disgust were more frequent than for other emotion pairs (with ps < .001), and the confusions between fear and surprise were themselves significantly more frequent than the confusions between anger and disgust, $t(35) = 5.9$, $p < .001$. However this greater confusability of fear and surprise than of anger and disgust held only when faces were recognized by participants from the same cultural background (Western Caucasian participants for Western Caucasian faces: $t(17) = 3.0$, $p < .05$; Western Caucasian participants for Chinese faces: $t(17) = 2.4$, $p > .1$; Chinese participants for Chinese faces: $t(17) = 6.9$, $p < .001$; Chinese participants for Western Caucasian faces: $t(17) = 0.6$, $p > .1$).

Our Chinese participants were all raised in China by Chinese parents, but they were all living in the UK at the time. We therefore used the data from the questionnaire concerning how long the Chinese participants had been in the UK to explore whether contact with Western Caucasian people might have influenced their performance to the Western Caucasian facial expressions. The time our Chinese participants had been in the UK ranged from 18 months to nine years and four months. To investigate whether contact with Western Caucasian people might have influenced the Chinese participants' performance with Western Caucasian expressions, we calculated the averaged similarity ratings for each set of faces for each of our Chinese participants, and then calculated the difference in similarity ratings between the two sets of faces (i.e. similarity ratings of Chinese faces minus those of Western Caucasian faces) and correlated these differences with time in the UK. From the social contact theories (Furl, Phillips, & O'Toole, 2002; Tanaka et al., 2004; Walker, Silvert, Hewstone, & Nobre, 2007) we might expect that the more time that Chinese participants have lived in a western country, the less would be the perceptual difference between the Western Caucasian and Chinese faces. However, our results (Figure 3.5A) were not consistent with this idea. Instead, they showed a significant positive relationship between rating differences and time spent in the UK, $r = 0.47$, $p = .05$; this result is in the opposite direction to the social contact hypothesis.

We also applied the same approach to the recognition accuracy data. A correlation analysis was also used to evaluate the relationship between each Chinese participant's time spent in the UK and their recognition difference between Chinese and Western Caucasian faces. The result showed a trend indicating that the longer the Chinese participants have been living in the UK, the less the identification difference between the Chinese and Western Caucasian faces. This is in line with the social contact hypothesis, but the trend did not reach a reliable level, $r = - 0.32$, $p = .20$ (Figure 3.5B).

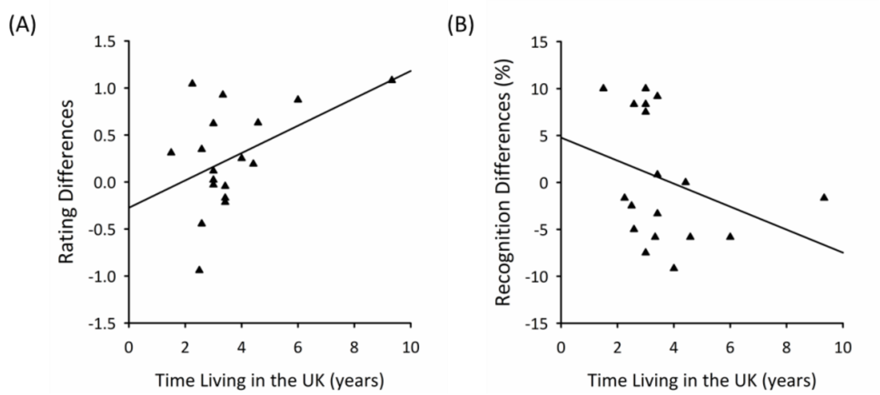


Figure 3.5 Scatterplots of Chinese participants' time living in the UK with their performance differences between Chinese and Western Caucasian faces in the perceptual similarity task (A) and the categorization task (B).

3.2.3 Discussion

In this experiment, we extended Yan et al.'s (2016a) study by investigating cross-cultural similarities and differences in perceiving and recognizing facial expressions of six basic emotions. We found a large amount of cross-cultural consistency in participants' perceptual similarity ratings of expression pairs, and also in the patterns of confusions from the categorization task.

Despite this general background of cross-cultural consistency, we found that a small own-group advantage for recognizing facial expressions is driven by the overall confusability of emotion categories. Our results only found a full cross-over interaction of participant group by face ethnicity for recognizing anger, some evidence of differences in recognition of disgust, and also a group difference between Western Caucasian and Chinese faces for

Western Caucasian participants. These results showed that the cross-cultural differences in expression processing were mainly centred on the recognition of anger and disgust.

Previous studies have shown that some pairs of facial expressions are more likely to be confused with each other; especially surprise with fear, and anger with disgust (Calvo & Lundqvist, 2008; Ekman & Friesen, 1976; Palermo & Coltheart, 2004; Wiggers, 1982). In our emotion categorization task, confusions among anger and disgust or fear and surprise were much higher than those of other expression pairs, and our two groups of participants showed a high consistency in the confusion patterns. However, as has been noted in other studies our participants made more confusion between fear and surprise expressions, compared with the confusions between anger and disgust, but despite this only anger and disgust recognition were linked to an own-group advantage. These results indicate that the own-group advantage in expression recognition cannot be explained simply by the degree of confusability of the expressions. We return later to the question of how it might therefore originate in our General Discussion.

In this experiment, we also investigated cross-cultural differences for surprise because Jack et al. (2012) reported that the surprise expression plays an important role in driving the own-group advantage in expression perception. This conclusion was linked by Jack et al. (2012) to a more general idea that East Asian participants rely considerably on the eye region and comparatively little on the mouth region in their mental representations of facial expressions. Although our findings from Experiment 1 did not lend support to the particular importance of surprise, we decided to further investigate Jack et al.'s (2012) more general position on the importance of the eye region in Experiment 2, by investigating whether there are cross-cultural differences in the holistic processing of facial expressions.

3.3 Experiment 2

We used the composite-expression paradigm devised by Calder et al. (2000) to investigate the holistic processing of own-race and other-race facial expressions by Western Caucasian and Chinese participants. From Jack et al.'s (2012) findings we predicted that if Chinese participants mainly use the eye region to internally represent facial expressions there should be a correspondingly reduced holistic processing of facial expressions. To test this prediction, we asked participants to identify facially expressed emotions from the upper (eye region) or lower (mouth region) parts of stimuli arranged in aligned composite (face-like) or misaligned (not face-like) formats.

3.3.1 Methods

3.3.1.1 Participants

Groups of 18 Chinese students brought up in mainland China with Chinese parents (13 females; mean age, 21.9 years) and 18 Western Caucasian students brought up in western countries with Western Caucasian parents (16 females; mean age, 21 years) were recruited from the University of York to participate in this experiment. All participants gave their written consent prior to the experiment and received a small payment or course credit. The University of York Department of Psychology Ethics Committee approved the study.

3.3.1.2 Stimuli

Based on our previous study (Yan et al., 2016a), we selected facial expressions of the three emotions that could be well-recognized from both the upper and lower part of the face, which are anger, fear, and happiness. The proportional recognition rates for these three emotions were 0.5, 0.6, and 0.9, respectively for the upper half faces, while the relative recognition rates for the lower part faces were 0.5, 0.7, and 0.9, respectively. Four exemplars of each emotion were selected from the stimuli used in Experiment 1. The stimuli were created by combining the upper and the lower halves of different facial expressions. This led in total to six possible upper/lower combinations; anger/fear, anger/happiness, fear/anger, fear/happiness, happiness/anger, and happiness/fear. The upper and lower halves of each stimulus were always taken from photographs posed by different models, because Calder et al. (2000) showed that the identities of the face parts had no effect on the holistic processing of facial expressions. All half faces were created by arbitrarily dividing each face through the middle of the bridge of the nose.

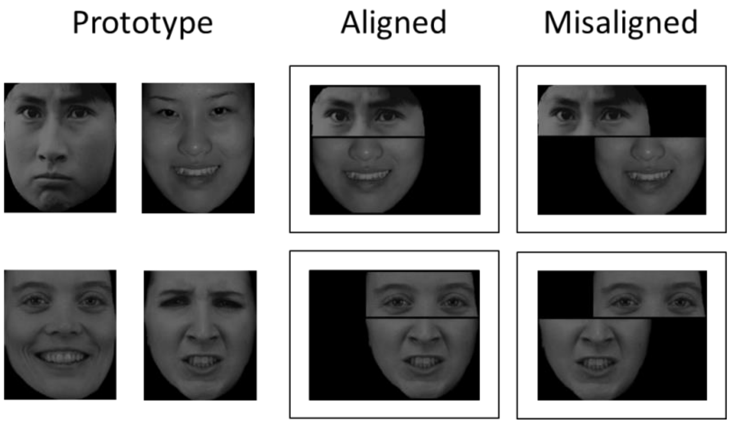


Figure 3.6 Examples of stimuli used in Experiment 2. The upper and lower half of different prototype expressions from one of the image sets (left) were combined to create aligned composite (middle) and misaligned (right) stimuli. The two prototype faces in the first row are Chinese models showing expressions of anger and happiness, respectively, from the Chinese Facial Affective Picture System (CFAPS; Wang & Luo, 2005; Gong et al., 2011) and the two prototype faces in the second row are Western Caucasian models showing happiness and anger expressions from the Karolinska Directed Emotional Faces (KDEF; Lundqvist et al., 1998).

Stimuli were presented in two different formats: aligned composites and misaligned images (Figure 3.6). The aligned expressions were presented in a face-like configuration, but (following the recommendation of Rossion & Retter, 2015) a narrow dark band was used to separate the upper and lower halves of each stimulus, so that participants could see that there were distinct top and bottom parts. The misaligned expressions were created from the same face parts as the aligned expressions, except that the upper and lower halves of the misaligned stimuli were misaligned horizontally. For these misaligned stimuli we followed Calder et al. (2000) by aligning the middle of the nose of the upper half faces with the edge of the lower half face. For half of the misaligned images, the upper half was shifted to the left side of the lower half, while for the other half of the misaligned stimuli the upper half was shifted to the right side of the lower half.

There were 4 stimuli for each of the 6 upper/lower expression combinations, giving a total of 24 aligned stimuli and 24 misaligned stimuli for each race set. When the misaligned faces were presented in the middle of the screen, neither the upper or the lower half faces was centralized in the screen. To match this, half of the aligned faces were presented in the same position as the left half of the misaligned faces and half in the same position as the right half of the misaligned faces (Figure 3.6). When viewed in the experiment the aligned images subtended a visual angle of approximately 8° x 7°, and the misaligned images were 8° x 10°.

3.3.1.3 Procedure

Participants viewed expression images using a computerized task programmed with PsychoPy software (www.psychopy.org). All participants made a three-alternative forced-choice (3AFC) involving judging the facial expression (anger, fear, or happiness) of the upper or lower half of both the Chinese and Western Caucasian faces. Responses were made via keypresses 1-3 for the expressions and the mapping between these emotion labels and response keys was counterbalanced across participants. The code for keypresses was always visible on screen. Each trial began with a central fixation cross for half a second, following which a stimulus was presented on the screen until the participant made a response. Participants were asked to respond as quickly and as accurately as possible.

All participants completed two blocks of trials. In one block, the task was to identify the facial expression of the upper half face, and in the other block the task was to identify the facial expression from the lower half. The sequence of these two blocks was counterbalanced across participants. The face stimuli for each block were identical, including 24 aligned and 24 misaligned Chinese faces and the same number of Western Caucasian faces. Faces of different races were presented in a block order of 'Chinese-Caucasian' or 'Caucasian-Chinese' which was counterbalanced across participants. Within each race set the 48 stimuli (aligned and misaligned images) were presented in a random order.

To ensure participants could correctly identify the upper or lower parts of the facial expressions, each block began with the presentation of only the half faces (upper or lower, as appropriate) that were used to create the aligned and misaligned stimuli. Participants were asked to identify the expression for each half face, and feedback was given in this part of the experiment only. The appropriate parts (upper or lower) of the 12 faces were each presented twice, making a total of 24 practice trials. After being familiarized with the half faces, no further feedback was given and the participants were also given 24 practice trials with the aligned and misaligned stimuli before the formal task in each block. These practice stimuli were made from the same part faces but with different combinations to those used in the main experimental trials.

3.3.2 Results

Our primary focus of interest is in reaction times for correct responses, with the expression composite effect being indexed by slower responses to aligned composite than to misaligned images. Slowing of responses to the aligned composites is thought to result from holistic perception of the face-like aligned expressions leading to a novel expression that interferes with identifying the expression in each face part (Calder et al., 2000). We conducted a mixed-ANOVA on the median correct reaction times (RTs) with Half (upper or lower part judgement), Face Ethnicity (Western Caucasian or Chinese faces), and Alignment (aligned or misaligned stimuli) as within-subject factors and Participant Group (Western Caucasian or Chinese participants) as a between-subject factor. This showed a significant main effect of stimulus Alignment, $F(1,34) = 29.3$, $p < .001$, partial $\eta^2 = 0.46$. Participants took longer to identify the parts of aligned expressions (1093ms) than misaligned expressions (968ms), consistent with the expression composite effect found in previous studies (Calder et al., 2000; Calder & Jansen, 2005; White, 2000).

There was also a significant Face Ethnicity by Participant Group interaction, $F(1,34) = 5.9$, $p < .05$, partial η2 = 0.15, indicating an own-group advantage in recognizing facial expressions (Figure 3.7). Further analyses showed that Chinese participants were faster at recognizing Chinese facial expressions than Western Caucasian expressions, $F(1,34) = 5.7$, $p < .05$, while there was no time difference between Western Caucasian and Chinese facial expressions for Western Caucasian participants, $F(1,34) = 1.1$, $p > .1$.

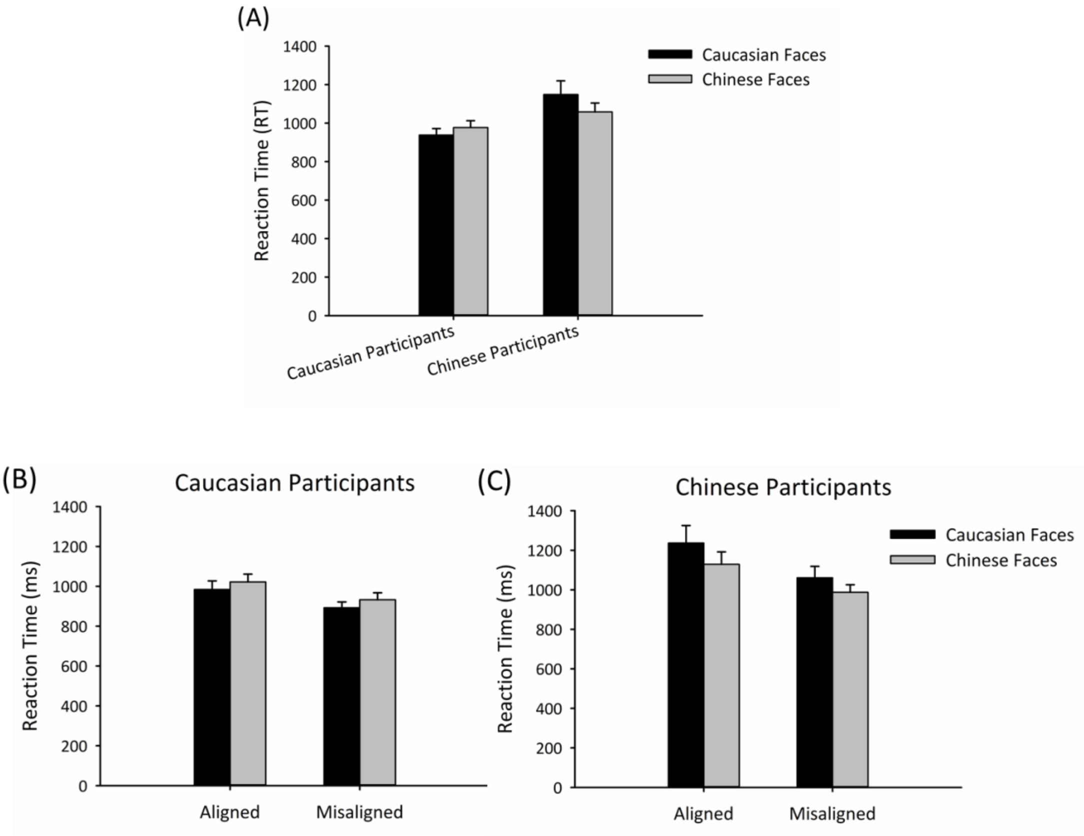


Figure 3.7 Correct reaction times for Western Caucasian and Chinese participants. (A) Overall correct reaction times (with standard error bars) for Western Caucasian and Chinese participants with the Western Caucasian and Chinese facial expressions in Experiment 2. (B) (C) Overall correct reaction times (with standard error bars) for Western Caucasian and Chinese participants recognising parts of aligned and misaligned stimuli created from upper and lower halves of Western Caucasian and Chinese expressions.

The ANOVA also found a significant main effect of face Half, $F(1,34) = 12.0$, $p < .001$, partial $\eta^2 = 0.26$, and this main effect was moderated by two two-way interactions; Face Ethnicity x Half, $F(1,34) = 6.1$, $p < .05$, partial $\eta^2 = 0.15$, and Alignment x Half, $F(1,34) = 6.2$, $p < .05$, partial $\eta^2 = 0.15$. Further analyses of these two-way interactions showed that the Chinese lower half faces were more quickly identified than the Western Caucasian lower half faces, while there was no difference for the upper half faces. None the less, for both the upper half and lower half faces, participants were always faster at recognizing facial expressions from misaligned than aligned faces. The main effect of Participant Group was also significant, $F(1,34) = 5.2$, $p < .05$, partial $\eta^2 = 0.13$, with Western Caucasian participants taking less time (958ms) than Chinese participants (1103ms) to identify the facial expression parts. No other significant effects were detected.

The most important RT findings, then, were a clear expression composite effect (main effect of Alignment) that was not modified either by Participant Group or by Face Ethnicity, indicating that the size of the expression composite effect was stable across participant and face ethnicities.

We also conducted an equivalent mixed-ANOVA on the arcsine transformed recognition accuracies. The results showed a significant main effect of Alignment, $F(1,34) = 39.7$, $p < .001$, partial $\eta^2 = 0.54$, indicating that participants were more accurate at recognizing facial expressions from misaligned stimuli versus aligned stimuli (see Figure 3.8) and demonstrating that there was not a speed-accuracy trade-off. The expression composite effect was again detected for both the upper half faces ($F(1,34) = 32.6$, $p < .001$) and the lower half faces ($F(1,34) = 3.3$, $p = .08$).

There were also significant main effects on accuracy for Face Ethnicity, $F(1,34) = 51.6$, $p < .001$, partial $\eta^2 = 0.60$, and Half, $F(1,34) = 16.8$, $p < .001$, partial $\eta^2 = 0.33$. Two significant two-way interactions were also detected: Face Ethnicity x Half, $F(1,34) = 32.4$, $p < .001$, partial $\eta^2 = 0.49$, and Half x Alignment, $F(1,34) = 11.5$, $p < .01$, partial $\eta^2 = 0.25$. Further analyses showed that the accuracy for recognizing Chinese face parts was higher than that of Western Caucasian face parts for only the lower half faces, $F(1,34) = 87.6$, $p < .001$. No other significant effects were revealed.

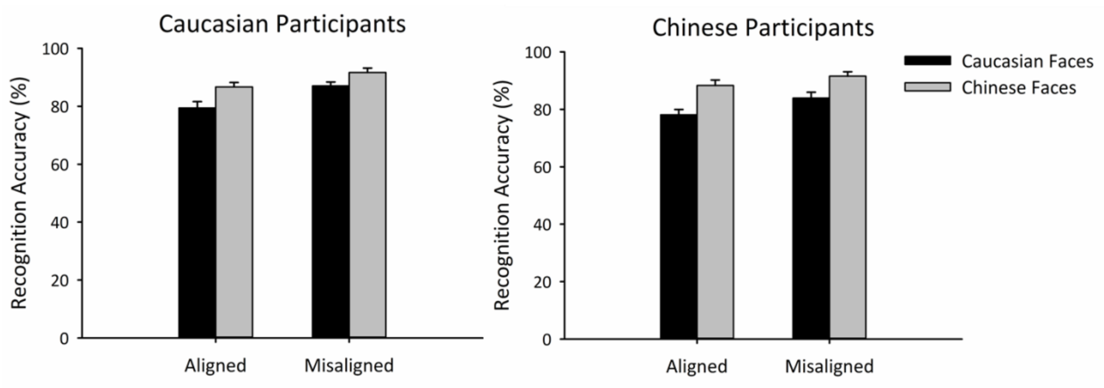


Figure 3.8 Overall recognition accuracies (with standard error bars) for Western Caucasian and Chinese participants with the Western Caucasian and Chinese aligned and misaligned stimuli in Experiment 2.

Because Experiment 1 only found own-group advantages for recognition of certain facial expressions (particularly anger), we carried out a supplementary analysis of the data from Experiment 2 to explore whether holistic processing was evident for each emotion category. We conducted a mixed ANOVA of the correct RTs which included Expression (anger, fear, and happiness) as an additional within-subjects factor. In order to examine the expression composite effect in each emotion category, we looked for significant effects involving the holistic processing of expressions. These were a main effect of Alignment ($F(1,34) = 33.8$, $p < .001$, partial $\eta^2 = 0.50$) and a three-way interaction of Half x Alignment x Expression

($F(2,68) = 4.9$, $p < .01$, partial $\eta^2 = 0.13$). The main effect of Alignment demonstrated an overall expression composite effect, in which participants needed more time to recognize facial expressions from aligned versus misaligned face parts.

In the three-way interaction of Half x Alignment x Expression, we found that participants recognized expressions faster from both the upper and lower parts of the misaligned anger and fear faces than from aligned faces, but the composite effect only existed when recognizing happiness from the upper part faces. This interaction therefore reflected the ease with which the smiling mouth is identified as a signal of happiness, leading to an absence of the expression composite effect for this condition only. No interactions involving Alignment or Participant Group were detected, indicating again that there were no group differences between Western Caucasian and Chinese participants in holistic processing of the three emotions. This was again inconsistent with the prediction based on Jack et al.'s (2012) study that Chinese participants would show reduced holistic processing for facial expressions.

3.3.3 **Discussion**

In Experiment 2, we used the composite effect to investigate holistic processing of facial expressions. We found a reliable expression composite effect; participants were faster and more accurate at recognizing facial expressions from half faces when they were in a misaligned arrangement that was not face-like. When the same half-faces were presented in a more face-like aligned composite format, responses to upper or lower parts were slowed and errors increased. These results indicated that facial expressions are processed in a holistic way. Importantly, this was true for both the Western Caucasian and Chinese participants, and for the Western Caucasian and Chinese expressions. The lack of cross-cultural differences in holistic perception of expressions is inconsistent with predictions based on Jack et al.'s (2012) view that Chinese participants focus on the eye region when internally representing facial expressions. Our results showed clearly that both groups of participants recognize facial expressions in a holistic way.

We did none the less find a small own-group advantage in overall reaction times, with Chinese participants spending less time recognizing Chinese faces than Western Caucasian participants, but no difference for Western Caucasian participants. However, this own-group advantage was not linked to differences in holistic processing of own-race versus other-race expressions. We also found equivalent holistic processing effects for each of the three facial expressions tested (with the minor exception of the lower parts of happy faces), and in both groups of participants.

To investigate the holistic processing of facial expressions, this experiment followed the design used by Calder et al. (2000), in which participants were required to recognize facial expressions from upper or lower half faces. As well as recognition tasks, a matching paradigm has also been widely used, in which participants are required to do a same/different matching judgement of two faces for each trial. However, there has been hot debate concerning choosing the optimal matching paradigm to capture face holistic processing ability (see Richler, Cheung, & Gauthier, 2011a, 2011b; and see opposite opinions in Rossion, 2013). However, much of this debate is irrelevant to recognition tasks, and for the first attempt to examine the holistic processing of own-race and other-race faces, this experiment only used the recognition task devised by Calder and colleagues (2000) to demonstrate holistic processing of facial expressions in Western Caucasian participants. Using the same task offers a good way of comparing between studies. In addition, the heated debate concerning the matching paradigm has centred on the fact that studies by Gauthier and colleagues (2011a, 2011b) don't use a misaligned face parts condition as a point of comparison to the aligned composite stimuli. Instead, I followed Rossion's (2013) recommendation by including both the misaligned and aligned conditions to eliminate influences of response competition (which are matched between aligned and misaligned conditions).

3.4 General Discussion

We investigated potential factors that might underlie cultural differences in facial expression recognition. In the first experiment, we replicated and extended Yan and colleagues' (2016a) results by showing that there was substantial cross-cultural consistency in perception of similarities between different pairs of expressions and in the patterns of confusion when categorizing expressions. The own-group advantage was only found in the categorization (not in the perception) of expressions, and mainly for expressions of anger and disgust. Even though we found more obvious categorization confusions between anger and disgust and also between fear and surprise than other expressions, which was consistent with the findings of previous studies (Calvo & Lundqvist, 2008; Ekman & Friesen, 1976; Palermo & Coltheart, 2004; Wiggers, 1982), only anger and disgust were linked to the own-group advantage. Therefore, the confusability of expressions cannot fully explain the own-group advantage in expression recognition.

In the second experiment, we explored another possible factor of engagement of holistic processing that might drive cross-cultural differences in expression recognition. We found a reliable expression composite effect for both groups of participants and both face ethnicities; participants were faster and more accurate at recognizing facial expressions from half parts of misaligned than aligned stimuli. These results indicate that for both the Western Caucasian and Chinese participants, expressions of both own-race and other-race faces are processed in a holistic way. This is inconsistent with the prediction based on Jack et al.'s (2012) hypothesis that Chinese participants mainly use the eye region to represent facial expressions. Moreover, since our results showed comparable magnitudes of holistic processing of expressions across Western Caucasian and Chinese participants, the own-group advantage in expression recognition cannot be explained by the engagement of holistic processing.

In both experiments, we none the less found a reliable own-group advantage in the overall recognition of facial expressions posed by own-race versus other-race members. However, this own-group advantage was small compared with the large amount of cross-cultural agreement, indicating that widely repeated claims that "they all look the same" overestimate the cross-cultural differences (Yan et al., 2016a, 2016b).

Even though we did not find group differences in holistic processing of facial expressions, some previous studies have linked the own-group advantages in the recognition of facial identity (rather than expression) to holistic processing, claiming a greater engagement of holistic processing by own-race than other-race faces (Tanaka et al., 2004; Michel et al., 2006a; Michel et al., 2006b). Alternatively, however, Hayward et al. (2013) have pointed to inconsistencies between previous findings involving the other-race effect for facial identity and argued that the key feature of own-race face advantages may lie in more effective processing of all types of face information (featural as well as holistic). Our study is the first to investigate potential cross-cultural differences in the holistic perception of facial expression and the discrepancy between our results for facial expression and these previous findings for facial identity processing is consistent with the idea that the underlying processing of facial expression and identity may be different (Bruce & Young, 1986; Calder & Young, 2005; Haxby, Hoffman, & Gobbini, 2000).

Since our results showed that the own-group advantage in facial expression recognition cannot be explained by either the confusability of emotions or the holistic perception of expressions, we can ask what then are the factors that cause the own-group advantages? One possible reason is that there are relatively minor cultural “stylistic” differences in the way in which certain emotions are expressed around a common overall template (Yan et al., 2016a), and we note two influences that may contribute to such differences for anger and disgust. First, compared to Western Western Caucasian individuals, people in Eastern Asian countries learn to avoid expressing negative emotions that might harm interpersonal and social harmony (Matsumoto, 1989; Matsumoto & Ekman, 1989). Second, and possibly linked to this, the meaning of disgust might be different across cultures (Han, Kollareth, &

Russell, 2015; Yoder, Widen, & Russell, 2016). Although Darwin (1872) and Rozin, Haidt and McCauley (1993) have argued that the evolutionary origins of disgust can be traced back to a rejection response to bad tastes and smells, other types of disgust can be added to this core disgust by 'an opportunistic accretion of new domains of elicitors, and new motivations, to a rejection system that is already in place'. These accretions can include responses to violations of moral or cultural rules and norms (Rozin et al., 1993). So there are clear possibilities for cultural differences. Compared with the Korean and Malayalam words for disgust, for example, Han et al. (2015) found that the English word disgust referred to more mixed emotional reactions to both physical and moral disgust scenarios. Similarly, by asking participants to choose an emotion label that best matched the emotion of several stories, Yoder et al. (2016) found that the facial expression that best described physical disgust stories was more like a ‘sick face’, while the more standard disgust facial expression and sometimes anger were more often chosen for the representation of moral violation stories. These findings coincide with our findings that own-group advantages were mainly evident for anger and disgust expressions, but not the more confusable expressions of fear and surprise.

In summary, the present study shows substantial cross-cultural consistency in perception of facial expressions of six basic emotions and also confusion patterns among emotions in Western Caucasian and Chinese participants. In contrast, cross-cultural differences in the categorization of expressions were real but small, and mainly existed for emotions of anger and disgust. Both Western Caucasian and Chinese participants process facial expressions in a holistic way and there were no differences in the engagement of holistic processing to own- and other-race faces. The own-group advantage in expression recognition cannot be explained by either the confusability of emotions or the holistic perception of expressions, but may reflect stylistic differences in the way that certain emotions are expressed within a common overall template.

4. Cross-cultural differences and similarities underlying other-race effects for facial identity and expression (Published in the Quarterly Journal of Experimental Psychology, 69, 1247−1254)

4.1 Introduction

The well-known other-race effect shows that cultural background can affect ability to recognise both face identity and facial expression. People are more accurate at recognising unfamiliar faces that seem to come from their own ethnic group (Meissner & Brigham, 2001; Chance & Goldstein, 1996; Brigham, Bennett, Meissner, & Mitchell, 2007). Similarly, an own-group advantage has also been found in facial expression recognition (Elfenbein & Ambady, 2002b; Jack, Caldara, & Schyns, 2012; Yan, Andrews, & Young, 2016a).

Although it is usually considered well-established that people are more accurate at recognising the faces and expressions of their own-group members, no studies have actually investigated the other-race effect in facial identity and expression at the same time. To date, substantial procedural differences between the tasks used to investigate identity and expression have precluded such a comparison, and widely-used methods also have significant limitations. For example, studies of identity recognition often use a recognition memory paradigm in which images of unfamiliar faces are studied and then tested for whether these learnt images can be distinguished from unstudied images. This task may in part tap face recognition abilities, but suffers the limitation that it also involves a substantial element of picture learning (Hay & Young, 1982; Longmore, Liu, & Young, 2008). On the other hand, studies of facial expression recognition usually use a forced-choice labelling paradigm that has been criticised as overestimating the degree of agreement (because expressions about which the participant is uncertain have to be assigned to the category forming the closest approximation) and also there may be problems in translating emotion labels (Matsumoto & Assar, 1992; Russell, 1994).

Here we test the other-race effect for both face identity and expression in tasks with equivalent structure that avoid the above pitfalls. We make use of adapted variants of a free-sorting task introduced by Jenkins, White, Montfort and Burton (2011). Their task involved giving participants twenty different images (everyday photographs) of two different unfamiliar faces, and asking participants to sort these into piles corresponding to different identities. Importantly, participants were not told that there were only two different faces in the set, so they were free to put together photos they perceived as showing the same face without any constraint.

For the present study, we adapted the Jenkins et al. (2011) task by creating sets of photographs showing 20 own-race or 20 other-race faces. These sets of 20 photos either comprised 5 varied images of each of 4 faces (identity sets) or 5 varied images of each of 4 emotional expressions (expression sets). Subject to these constraints, there was no attempt to constrain the different images of each identity or each expression so that they would particularly resemble each other, in line with Jenkins et al.'s (2011) 'ambient images' approach. Participants were then asked to sort the 20 images in each identity set into piles in which they perceived each face as having the same identity, and the 20 images in each expression set into piles in which they perceived each face as having the same expression. In this way, we created identity and expression tasks with equivalent demands ("sort the photographs into piles"). No verbal labels or categories (other than the requirement to sort by identity or by expression), and no fixed forced-choice requirement (participants were free to create as many or as few piles as they thought appropriate).

To ensure that any cross-cultural differences were not simply due to the images themselves, we used a full crossover design in which participants from Chinese and Western Caucasian backgrounds sorted both Chinese and Western Caucasian faces. We predicted that Chinese participants would make more confusions for Western Caucasian faces, while Western Caucasian participants would make more confusions for Chinese faces.

This novel procedure allowed us to address a key question concerning the magnitude of cultural differences that are reflected in other-race effects. Many research studies create the impression that the underlying cultural differences are large, as reflected in everyday opinions such as “they all look the same” (Feingold, 1914; Vizioli, Rousselet, & Caldara, 2010). However, in a recent study that investigated cultural differences between Chinese and British participants with very different methods involving perceptual similarity ratings and forced-choice categorization, we found that the other-race effect in forced-choice expression recognition was quite small (5%-9%) in comparison to the level of cross-cultural agreement (Yan, et al., 2016a). Here, we use the free-sorting procedure to determine the extent of cross-cultural agreement and differences by correlating the patterns of response made by Chinese and Western Caucasian participants, offering a complementary perspective on Yan et al.'s (2016a) findings.

4.2 Methods

4.2.1 Participants

Twenty Chinese students brought up in mainland China with Chinese parents (mean age, 22.6 years) and 20 Western Caucasian students brought up in western countries with Western Caucasian parents (mean age, 20.1 years) were recruited from the University of York. None of the participants were familiar with any of the stimulus faces. All participants gave their written consent prior to the experiment and received a small payment or course credit. The University of York Department of Psychology Ethics Committee approved the study.

4.2.2 Stimuli

Two sets of 20 Western Caucasian and two sets of 20 Chinese faces were created for the identity sorting task, and two sets of 20 Western Caucasian and two sets of 20 Chinese faces for the expression sorting task.

For the identity task, each set contained five images of each of 4 male Australian or 4 male Chinese celebrities selected and downloaded from the internet (20 images per set). To ensure that these faces were unfamiliar to participants, we chose Australian celebrities we thought unlikely to be known to our Western Caucasian (mostly British) participants, and Chinese celebrities from Taiwan and Hong Kong who would not be known to participants from mainland China. Participants who recognised any of the faces were replaced. To select the specific photographs used, we followed the criteria adopted by Jenkins et al. (2011): (1)

exceeding 150 pixels in height, (2) showing faces from an approximately frontal viewpoint, (3) free from occlusions.

For the expression task, we used stimuli from sets previously used by Yan, et al. (2016a); the Chinese Facial Affective Picture System (CFAPS) (Wang & Luo, 2005; Gong, Huang, Wang, & Luo, 2011) posed by Chinese models, and the Karolinska Directed Emotional Faces (KDEF) (Lundqvist, Flykt, & Öhman, 1998) posed by Western Caucasian models. These sets were chosen because the instructions given to the models were simply to pose expressions as best as they could, without specific requirements concerning which facial muscles to move, leading to variability in how the expressions were posed. Each set contained five randomly selected images of each of 4 negative expressions (anger, disgust, fear, and sadness). All images were converted into greyscale and printed onto laminated cards extending 38 mm in width and 50 mm in height.

4.2.3 Procedure

Each participant was asked to complete the sorting task for the 8 different sets of 20 stimuli; 2 Chinese Identity sets, 2 Western Caucasian Identity sets, 2 Chinese Expression sets, and 2 Western Caucasian Expression sets. Participants were given a shuffled deck of 20 face images (one of the eight sets). Their task was to sort the images into piles according to the identity or expression of the face, with images of the same person (in the identity task) or the same facial expression (in the expression task) grouped together into one pile. No other information was given to participants, so they could create as many piles and put as many images into each pile as they wished. The order of the identity and expression sorting tasks and the face sets were counterbalanced between participants. There was no time limit in each task, but most participants took about half an hour in total to complete sorting all 8 sets.

4.3 Results

As an initial evaluation of other-race effects for identity and expression, three dependent variables were recorded for each set; the number of piles created (i.e. the number of categories a participant thought there were for each set of stimuli), confusions (i.e. the number of faces from different categories that were grouped into the same pile) and the time taken to achieve the sorting. The notion that “they all look the same” is most clearly captured by confusions in which different people are mistaken for the same person. Following Jenkins et al. (2011), confusions were calculated by subtracting 1 from the number of categories represented in each pile; so a score of zero would indicate that only one identity or emotional expression was present in a pile, a score of 1 for two categories in the same pile, and so on. These individual pile scores were then summed to create an overall confusion score for each stimulus set. Note that in the sorting task there are actually two types of possible error. One type of error is that two different people are seen as only one, and this type of error can be captured by “confusions”. The other type of error is that the same person is seen as two different ones, and this can be captured by the “number of piles”. These two measures were previously used by Jenkins and colleagues (2011), and by starting with these measures we could compare our data to previous findings. Megreya and

Burton (2006) have shown that in a similar same/different matching task, the equivalent different types of matching errors do not correlate.

Performance for the two sets used for each sorting task (Chinese Identity, Western Caucasian Identity, Chinese Expression and Western Caucasian Expression) was then averaged for each participant.

A three-way ANOVA was conducted for each of these three measures (confusions, number of piles, sorting time) with Face Ethnicity (Chinese faces, Western Caucasian faces) and Task (Identity, Expression) as within-subject variables, and Participant Group (Chinese participants, Western Caucasian participants) as a between-subject variable.

Figure 4.1 shows the key measure of number of confusions for the identity and expression tasks. The ANOVA revealed no main effect of Task, Face Ethnicity or Participant Group on the number of confusions. However, there was a significant interaction between Face Ethnicity and Participant Group ($F(1,38) = 37.86$, $p < .001$, partial $\eta^2 =0.50$). Further simple effects analysis showed that there were significant differences in the confusion made by Chinese participants between Chinese and Western Caucasian faces ($F(1,38) = 12.94$, $p < .001$), and in the confusion made by Western Caucasian participants between Chinese and Western Caucasian faces ($F(1,38) = 26.06$, $p < .001$), indicating the existence of a classic other-race effect with a crossover interaction. This interaction was not qualified by any three-way interaction of Face Ethnicity × Task × Participant Group, $F(1,38) = 0.71$, $p = .4$, partial $\eta^2 = 0.02$, indicating that the underlying pattern of a crossover other-race effect was not affected by the task. There was also an unexpected significant interaction of Task × Participant Group, $F(1,38) = 5.22$, $p < .05$, partial $\eta^2 = 0.12$, reflecting a borderline difference between the number of confusions made by Chinese than by Western Caucasian participants in the expression task, $F(1,38) = 3.96$, $p = .05$. No other significant effects were found.

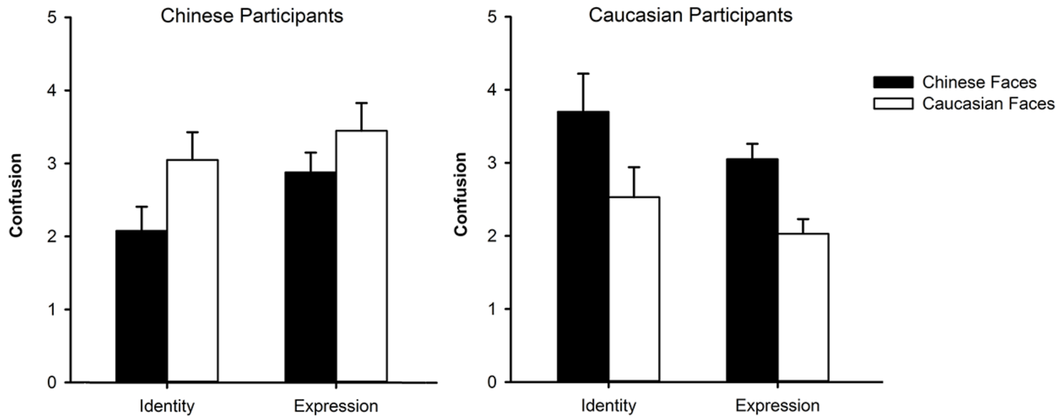


Figure 4.1 Mean confusions (derived from piles containing more than one identity or more than one emotional expression) for Chinese and Western Caucasian participants in facial identity and expression sorting tasks involving Chinese and Western Caucasian faces (with standard error bars).

Figure 4.2 shows the number of piles created for the identity and expression tasks. The 3-way ANOVA showed a significant main effect of Task, with participants making more piles on the identity task compared to the expression task ($F(1,38) = 58.13$, $p < .001$, partial $\eta^2 = 0.61$). There was no effect of Face Ethnicity ($F(1,38) = 3.20$, $p = .08$, partial $\eta^2 = 0.08$) or Participant Group ($F(1,38) = 0.78$, $p > .1$, partial $\eta^2 = 0.02$). However, there was a significant interaction between Face Ethnicity and Participant Group ($F(1,38) = 9.26$, $p < .01$, partial $\eta^2 = 0.20$). This was because there was a significant difference in the number of piles made by

Western Caucasian participants for Chinese faces compared to Western Caucasian faces (F(1,38) = 11.68, p < .01), whereas no reliable difference was observed between the number of piles made for Western Caucasian and Asian faces by Chinese participants (F(1,38) = 0.78, p > .01). There were no other significant effects.

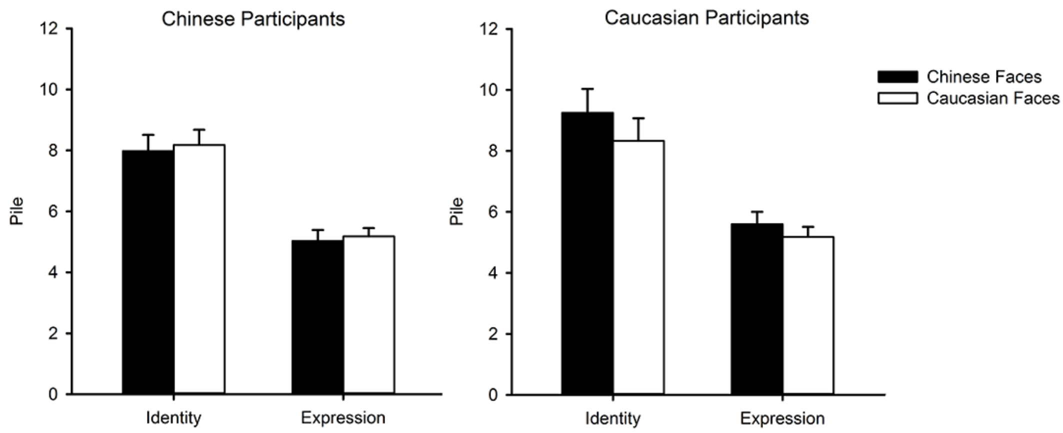


Figure 4.2 Mean numbers of piles created by Chinese and Western Caucasian participants in facial identity and expression sorting tasks involving Chinese and Western Caucasian faces (with standard error bars).

Figure 4.3 shows the sorting time for the identity and expression tasks. The ANOVA found a significant main effect of Face (F(1,38) = 4.7, p < .05, partial η^2 = 0.11), with slightly more time spent on the Chinese faces compared with the Western Caucasian faces (Chinese faces: 3.9 min; Western Caucasian faces: 3.5 min). There was also a significant main effect of Task (F(1,38) = 63.73, p < .001, partial η^2 = 0.63), with more time spent on the identity task than the expression task (Identity task: 4.7 min; Expression task: 2.7 min). No other effects reached significance.

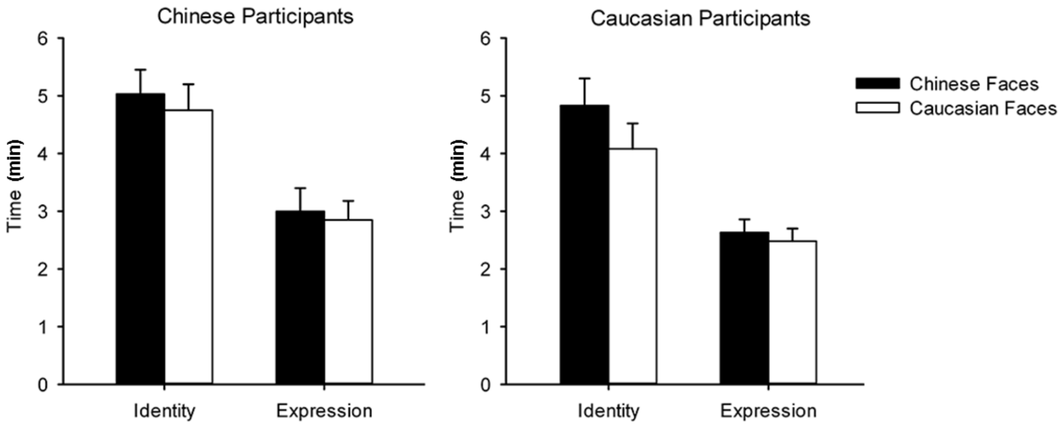


Figure 4.3 Mean sorting time (in minutes) for sets of 20 stimuli by Chinese and Western Caucasian participants in facial identity and expression sorting tasks involving Chinese and Western Caucasian faces (with standard error bars).

The main finding from these analyses, then, was the Face Ethnicity x Participant Group interaction for confusions shown in Figure 4.1. Next, we asked whether the pattern of responses was similar or different across the two groups of participants. To do this, we generated the full response matrix for each stimulus set for each group of participants. Each cell in a response matrix indicated the number of times that participants sorted two different images into the same pile. Figure 4.4 shows examples of the response matrices for the groups of participants in one Western Caucasian identity and one Chinese expression sorting task.

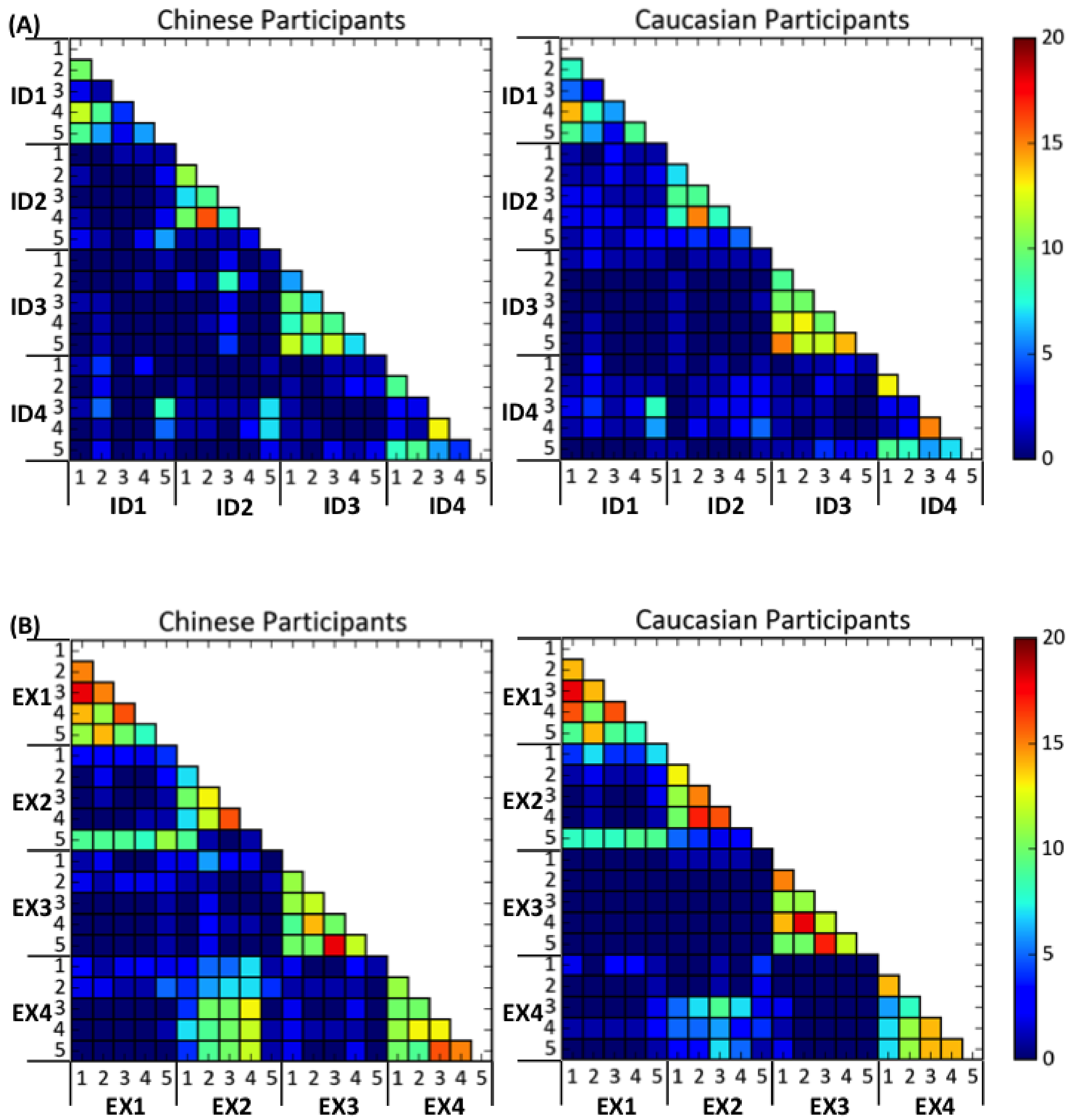


Figure 4.4 Response matrices for Chinese and Western Caucasian participants for one Western Caucasian Identity set (A) and one Chinese Expression set (B). The X- and Y-axes indicate the 5 different images of each of 4 identities/expressions. Each cell in the matrix represents the number of times that two images were sorted into the same pile by participants in the group. Different images that are seen as the same person or as expressing the same emotion will thus show up as more brightly coloured, and an idealised solution in which every identity/expression is seen as intended would lead to a set of bright

regions involving right-angled triangles along the diagonal with opposite and adjacent sides that are 4 cells long. The correlations of the response matrices between Chinese and Western Caucasian participants in both cases were 0.90, *ps* < .001.

From these response matrices we calculated a measure of cross-cultural agreement based on the overall correlations between the response matrices of Chinese and Western Caucasian participants for all 8 sets of stimuli. The importance of this correlation-based measure is that it incorporates both the extent of cross-cultural agreement and differences within a common overall metric. The r value among the two groups never fell below 0.70, and could rise as high as 0.91, as shown in Table 4.1. Strikingly, even though the ANOVA found a reliable other-race effect for both groups of participants, their sorting solutions none the less showed high consistency across cultures.

Table 4.1 Correlations between the sorting solutions of Chinese and Western Caucasian participants for the eight different sorting tasks. The Overall correlations use all the data from the corresponding response matrices (as shown in Figure 4.4). The Within correlations use only those cells in each matrix where responses should be assigned to the same category (for example, where two different images show the same identity or the same expression), and the Between correlations involve the remaining cells where the stimuli come from different categories (i.e. where two different images show different identities or different expressions). Significant correlations (*ps* < .001) were obtained for each measure, indicating a compelling pattern of agreement across cultures.

		Chinese Identity	Caucasian Identity	Chinese Expression	Caucasian Expression
Overall	Set1	0.70	0.85	0.90	0.90
	Set 2	0.84	0.90	0.80	0.91
Within	Set1	0.81	0.83	0.81	0.80
	Set 2	0.74	0.86	0.57	0.70
Between	Set1	0.51	0.47	0.75	0.73
	Set 2	0.36	0.57	0.61	0.62

However, these high correlations might be driven simply by agreement over the most clear cases in which stimuli were assigned to the same category. In Figure 4.4, an idealised solution in which every identity/expression is seen as intended would lead to a set of bright regions involving right-angled triangles along the diagonal with opposite and adjacent sides that are 4 cells long. We therefore also correlated the response patterns separately for these triangular within-category regions and the remaining between-category regions, as shown in Table 1. Substantial correlations (identity task: $r = 0.64 \pm 0.19$, expression task: $r = 0.69 \pm 0.09$) were still obtained, indicating a compelling pattern of agreement across cultures.

4.4 Discussion

We report the first systematic study of cultural differences in both facial identity and facial expression recognition. With a novel paradigm that matched the task demands of identity and expression recognition and avoided constrained forced-choice or verbal labelling requirements, we demonstrated other-race effects of comparable magnitude across the identity and expression tasks. Western Caucasian participants made more confusions for the identities and expressions of Chinese than Western Caucasian faces, while Chinese participants made more confusions for the identities and expressions of Western Caucasian than Chinese faces.

Although our paradigm matched task demands, participants created more piles and took longer to sort identities than expressions, suggesting a difference in overall task difficulty. None the less, a full crossover interaction between Face Ethnicity and Participant Group was evident for the confusions. The crossover interaction was not evident for the numbers of piles created. At present, we do not have an account as to why one measure should be more informative than the other, and it is clear that the measures may not be independent (for example, creating more piles may reduce the number of potential confusions). However, our data also allow us to measure the extent of cross-cultural similarities in the patterns of response, using a measure that combines information about both piles and confusions. By correlating the response matrices across Chinese and Western Caucasian participants, we showed that there is actually a considerable amount of cross-cultural agreement. For our 8 sets of stimuli, the overall cross-cultural correlation between Chinese and Western Caucasian participants' patterns of response never fell below 0.70, and could rise as high as 0.91. Both groups of participants even showed high consistency of their response patterns for images that fell in the same or different identity/expression categories. Consistent with our previous finding (Yan, et al., 2016a), this present study also provided evidence showing substantial cross-cultural agreement. The idea that other-race faces all (or even mostly) look the same is clearly overstated.

An interesting point is that we found the other-race effect for sorting simultaneously presented unfamiliar face identities. Most studies of the other-race effect in identity recognition have been based on recognition memory tasks, but recent studies have also found evidence of the other-race effect at the perceptual level. For example, in a task where participants were required to find a target face in a line-up of 10 faces, Megreya and colleagues (2011) found that both British and Egyptian participants were worse at matching other-group faces than own-group faces. Our results add evidence to confirm that difficulty in perceptual encoding of unfamiliar faces contributes to the other-race effect.

This study found a similar other-race effect in both the perception of facial identity and expression, indicating some common mechanisms in the cross-cultural processing across these two domains. However, there has been debate cornering the relationships between these two aspects of processing (Bruce & Young, 1986; Calder & Young, 2000; Ganel & Goshen-Gottstein, 2004; Schweinberger & Soukup, 1998). This issue is discussed in more detail in Section 6.5.

To summarise, our findings demonstrated the other-race effect across facial identity and expression with equivalently-structured tasks. However, the opinion that these cross-cultural differences are large was rejected as we found a substantial amount of cross-cultural agreement in both identity and expression processing.

5. The neural representation of facial identity and expression of own- and other-race faces

5.1 Introduction

The previous chapters in this thesis have explored the cross-cultural processing of own-race and other-race facial identity and expression in Western Caucasian and Chinese participants. Reliable advantages for the recognition of own-race versus other-race faces were found for facial identity and expression. However, the results also showed that the magnitude of the other-race effect was quite small compared to the large amount of cross-cultural agreement. This chapter explores the neural representations underlying the cross-cultural differences in the human brain.

Functional neuroimaging studies have shown that certain areas in the brain response particularly to faces rather than other objects, including the OFA (Occipital Face Area), FFA (Fusiform Face Area), and posterior STS (Superior Temporal Sulcus) (Kanwisher et al., 1997; Haxby, et al., 2000). Haxby and Colleagues (2000) argue that in the core system of face processing model, the OFA is involved in the structural-encoding phase of face processing, and then inputs information to both the FFA and the STS regions for further processing. The route to the FFA is involved in the encoding of invariant characteristics of faces such as identity, while the route to the posterior STS is involved in the processing the changeable aspects of faces, such as facial expressions.

Previous studies have found differences in the response to own-race and other-race faces in face selective regions, such as the FFA. For example, Golby et al. (2001) found that the

FFA responded more strongly to own-race faces versus other-race faces for both European-American and African-American participants. However, further analyses found that performance in the behavioral task, that is the score difference in face memories between own- and other-race faces, was only positively correlated with the activations in the left FFA. Feng et al. (2011) investigated the neural correlates of race categorization of own- and other-race faces. Their results found greater neural responses when categorizing own-race compared to other-race faces in the FFA, but, similar to Golby and colleagues (2001), only the neural activation in the left FFA was correlated with the behavioural categorization performance. Kim and colleagues (2006) also showed a larger response to own-race compared to other-race faces in the FFA, but this was only apparent for unfamiliar faces. Kim et al. (2006) suggested that this was driven by the long-term differences in the perceptual experiences to own-race compared to other-race unfamiliar faces, but the race differences could be overridden by the general familiarity effect. Chiao and colleagues (2008) explored the neural mechanisms underlying the facial expression (but only fear) recognition between different face ethnicities. Their results showed that both Japanese and Western Caucasian participants showed greater amygdala activation to their own- than other-race facial expressions, reflecting the own-group advantage in the processing of own-race expressions. However, Chiao et al. (2008) only investigated the neural responses in the amygdala, so it was not clear whether the own-group advantage, especially in facial expression processing, also existed in other face-selective regions, such as the FFA and the STS.

Using multivoxel pattern analysis (MVPA), Brosch and colleagues (2013) showed that the activation patterns in the FFA to white and black faces could be predicted based on the BOLD patterns obtained from the training set, but this only applied to participants with high pro-White bias. This implies that the differences in FFA activation to different race faces are also modulated by other factors, for example, people's attitude to other-race faces. Similarly by using the MVPA method, Natu, Raboy and O'Toole (2011) found a reliable own-race bias in the broader ventral temporal (VT) cortex, rather than the FFA alone. Further analyses however showed that this cross-cultural difference in the activation patterns was time sensitive. The temporal own-race bias was found in the first few time points of the block, for both the FFA and the broader VT cortex, but it attenuated rapidly.

This chapter aims to systematically investigate the neural mechanisms of the ORE in face perception across facial identity and expression in Haxby et al.'s (2000) core face-selective regions. The predictions are that these three regions will show high sensitivity to face stimuli, and they will show differences in the responses to own-race versus other-race faces.

To understand the neural correlates of the ORE, we used the powerful fMRI-adaptation technique (Grill-Spector, Henson, & Martin, 2006). The principle behind fMRI adaptation is that repetition of a stimulus causes a reduction or habituation in the neural response, which leads to a lower fMRI signal. The sensitivity of the neural representation can then be determined for different changes to the stimulus. If the underlying neural representation is insensitive to a particular type of change in the stimulus, the reduction in fMRI signal for this type of change will be similar to the overall reduction produced by repetitions of identical stimuli. However, if the underlying neural representation is sensitive to this change, the fMRI signal will remain at its original (non-adapted) level. In the present study, we compared adaptation to identity and expression in own-race and other-race faces. Previous studies have found adaptation to identity in the FFA and expression in the STS (Andrews & Ewbank, 2004, Andrews et al., 2010; Harris et al., 2012; 2014b). Our hypothesis is that these effects should be larger for own-race faces compared to other-race faces.

Vizioli et al. (2010) investigated fMR-adaptation or repetition suppression to own-race versus other-race faces in an ERP study. In their experiment, the first adaptor stimulus was followed by a test face of either the same or different identity (only facial identity was tested) of the same race. The results found significant suppressions in N170 (a negative deflection in the Event-Related Potential signal occurring roughly 170 ms after stimulus onset, peaking at occipitotemporal sites, see Rossion & Jacques, 2008 for a review) to the same identity of only the same-race faces in both East Asian and Western Caucasian participants, while similar suppression responses were not found regardless of changes in facial identity for other-race faces, providing a neurophysiological correlate of the "they all look alike" perceptual experience. However, this result is inconsistent with the findings of previous chapters in this thesis that show significant cross-cultural similarity.

This chapter has two aims: (1) to define regions that are sensitive to facial identity and expression and (2) to investigate the neural representations of own-race and other-race faces in the core face-selective regions. To address these questions, the block-design fMR-adaptation method is used by showing participants own- and other-race facial images with either the same or different combinations of facial identity and expression in different blocks. The hypotheses are: (1) brain regions such as the FFA will show a significant adaptation to facial identity; and (2) there will be small but reliable differences in the responses to own- and other-race faces in the three brain regions.

5.2 Methods

5.2.1 Participants

Fourteen Western Caucasian (10 females; mean age, 22 years old) and 14 Chinese (11 females; mean age, 23 years old) participants were recruited in the study. All participants were right handed, had normal or corrected to normal visual and no history of mental illness. All participants gave their written consent and the study was approved by the York Neuroimaging Centre Ethics Committee.

5.2.2 Stimuli

For the experimental scan, East Asian and Western Caucasian faces were selected from the Montreal Set of facial displays of emotion (MSFDE) with 6 identities and 6 expressions (anger, disgust, fear, happiness, neutral, and sadness) of each race category (Beaupré & Hess, 2005). All images were presented in grey scale and were 280 pixels high and 224 pixels wide. When viewed from 57cm away, each image extended approximately 7.5° high and 6° wide. The stimuli used here were not from the same sources as those used from previous chapters because that there was no identity information about the faces in the original Chinese faces set, which thus did not meet our need to creates conditions with the same or different identities. In addition, the East Asian and Western Caucasian faces used in the scanning were now from the same face set, which could remove some irrelevant noise variation from the images themselves (mainly differences in the physical properties between face stimuli).

For the localiser scan, different stimuli from those used in the experimental scan were used. The face stimuli were selected from face sets used in the previous three chapters; the Chinese Facial Affective Picture System (CFAPS) (Wang & Luo, 2005; Gong, Huang, Wang, & Luo, 2011) posed by Chinese models, and the Karolinska Directed Emotional Faces (KDEF) (Lundqvist, Flykt, & Öhman, 1998) posed by Western Caucasian models. Four images of each of 5 expressions (anger, disgust, fear, happiness, and sadness) were randomly selected from these sets. Scene images were selected from images used in previous studies (Watson, Hartley & Andrews, 2014; Watson, Hymers, Hartley & Andrews, 2016). All images were presented in grey scale and were 256 pixels high and 256 pixels wide, and when viewed from 57cm away, each image extended approximately 7° high and 7° wide.

5.2.3 fMRI Experimental design

Participants viewed images from four stimulus conditions with two races of faces in different blocks (resulting in a total of 8 blocks): (1) same identity same expression (sIsE), (2) same identity different expression (sIdE), (3) different identity same expression (dIsE), (4) different identity different expression (dIdE). Examples of images in these conditions are shown in Figure 5.1. In the experiment, each stimulus condition was presented in a block design consisting of 6 images, presented at one per 800ms with a 200 ms inter-stimulus interval. The interval between the blocks was 9 seconds. Each condition was repeated 6 times in a counterbalanced order, resulting in a total of 48 blocks and a scan time of 12 min. All participants completed the task with the same sequence. The task of the participants was to monitor the presence of a red dot that was superimposed on one image in each block, in order to ensure they were attending to the stimuli. Once they noticed the red dot, they were instructed to press a button as soon as possible. The red dot could show up in any location on the image, and was counterbalanced across conditions.

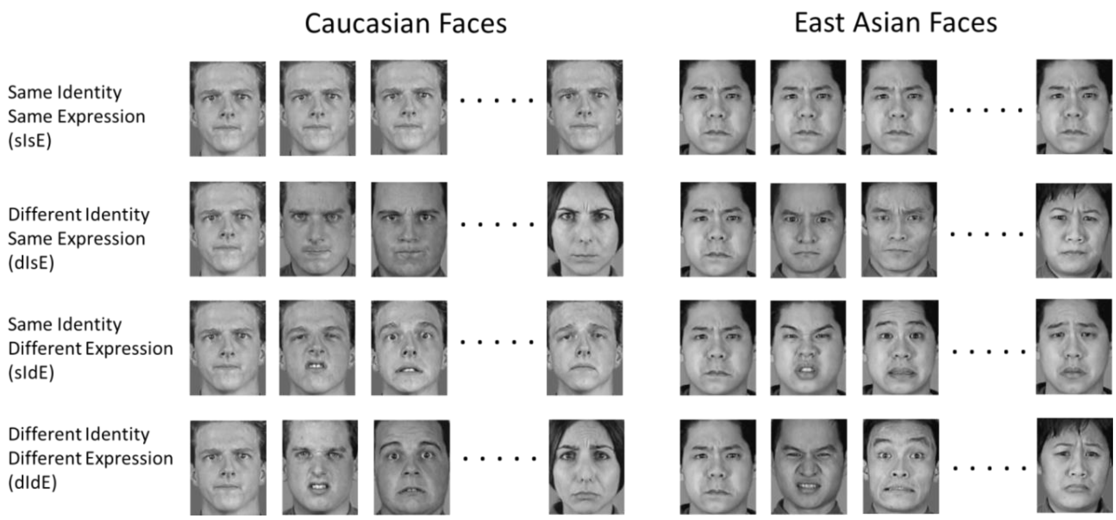


Figure 5.1 Examples of stimulus conditions represented with Western Caucasian and East Asian faces in the experimental scan.

A separate localizer scan was conducted after the experimental scan to identify regions in each participant's brain that responded more to faces versus non-faces. Each participant viewed blocks containing images taken from one of three different categories: (1) faces (East Asian and Western Caucasian faces showing different identities and different expressions), (2) scenes, or (3) scrambled faces created by phase scrambling the face stimuli. Examples of images in these conditions are shown in Figure 5.2. Images from different conditions were also presented in a block design consisting of 6 images, with each image presented for 800 ms with a 200-ms inter-stimulus interval. The interval between blocks was 9 seconds. Each condition was repeated 5 times, giving a total of 15 blocks and a scan time of about 4 min. The task of the participant was again to detect a red dot imposed on 20% of images and respond by pushing a button.

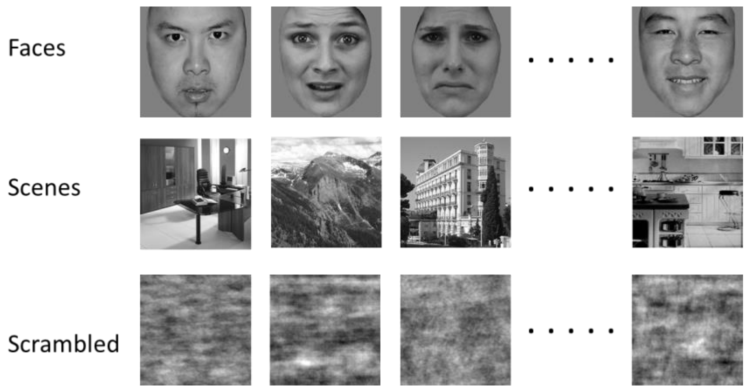


Figure 5.2 Examples of stimulus conditions in the localiser scan.

5.2.4 Imaging Parameters

All fMRI scans were carried out using a GE 3 Tesla HD Excite MRI scanner at the York Neuroimaging centre (YNiC) at the University of York. Experimental data were collected from 240 volumes each containing 38 axial slices via a gradient-echo EPI sequence (TR = 3s, TE = 32.7 ms, flip angle = 90°, FOV = 28.8 x 28.8 cm, matrix size = 128 x 128, voxel dimensions = 2.25 x 2.25 mm, slice thickness = 3 mm). Visual stimuli were back-projected onto an in-bore screen at a distance of approximately 57 cm from participants. For each participant, a T1-weighted structural MRI, a gradient-echo EPI, and a high resolution T1

Flair acquired in the same orientation planes as the fMRI protocol were acquired. While for the localiser scan, data was collected from 75 volumes each containing 38 axial slices via a gradient-echo EPI sequence.

5.2.5 fMRI analysis

Univariate analysis of the fMRI data was performed with FEAT v6 (http://www.fmrib.ox.ac.uk/fsl). For each scan, the initial 9 s of data was removed to reduce the effect of magnetic stimulation. Motion correction (MCFLIRT) was applied followed by temporal high-pass filtering (cutoff, 0.01Hz). Spatial smoothing (Gaussian) was applied at 6mm (FWHM). Individual data were entered into a group-level analysis using a mixed-effects design (FLAME, http://www.fmrib.ox.ac.uk/fsl). Face-selective regions were defined by the average statistical maps of faces > scrambled faces and faces > scenes for each individual in the localiser scan. A flood-fill algorithm was used to generate regions of interest (ROIs) with a size of 50 voxels each by adjusting the threshold iteratively. Only ROIs with voxels above a threshold of $Z = 2.3$ were included in the analysis.

All voxels in a given ROI region were averaged together to create a single time series for each participant, with the units of image intensity from the time series plot converted to percentage signal change. The peak response to each condition was taken at TR2, corresponding to 6 s after stimulus presentation.

The individually-defined ROIs method has been widely used in many studies (Flack et al., 2015; Harris et al., 2012, 2014b; Weibert & Andrews, 2015). The advantage of using this method is that it takes into account inter-individual variation in the functional locations of

ROIs. However, it is sometimes difficult to find ROIs in all individuals, which can cause issues such as the assumption of random and independent sampling being violated if these samples are excluded from further analyses (Swallow et al., 2003). Also, since the locations of ROIs can be slightly different owing to individual differences, statistical comparisons will be based on responses extracted from spatially different regions even though these may be closer to functionally equivalent. To create a complementary analysis that can avoid these potential issues induced by the individually-defined ROIs, group-level defined ROIs were also generated according to the average statistical maps of group faces > scrambled faces and faces > scenes across all the participants. The flood-fill algorithm was used again, but this time to achieve ROIs with a size of 500 voxels each by adjusting the threshold ($z > 2.3$) iteratively. This method, however, cannot account for individual functional variability as different voxels could be included or excluded mistakenly in each participant. The featquery analysis (http://www.fmrib.ox.ac.uk/fsl) was then conducted with the group-level masks to extract the mean values of responses within that certain ROI for each participant. The mean responses were then entered into mixed-ANOVAs to determine significant differences between conditions.

Finally, a whole brain analysis was conducted to determine whether other regions might demonstrate an adaptation effect to facial identity and expression. For each race of face for each group of participant, two comparisons were made (sIdE – sIsE, dIsE – sIsE), resulting in a total of 8 comparisons. The resulting statistical maps for each individual were combined using a higher-level mixed effect analysis (FLAME, FSL). The combined statistical maps were threshold at $Z > 3.0$, $p < 0.05$ (cluster corrected).

5.3 Results

The localizer scan reliably identified two face-selective regions: the FFA and the OFA (Figure 5.3). The right posterior STS and the left and right amygdala were only identified in a few participants and were therefore not included in the analysis. MNI coordinates of the peak voxels for each ROI in each hemisphere and corresponding thresholds are given in table 5.1. The peak voxels have similar coordinates to those found in previous studies (Harris et al., 2012, 2014b; Weibert & Andrews, 2015).

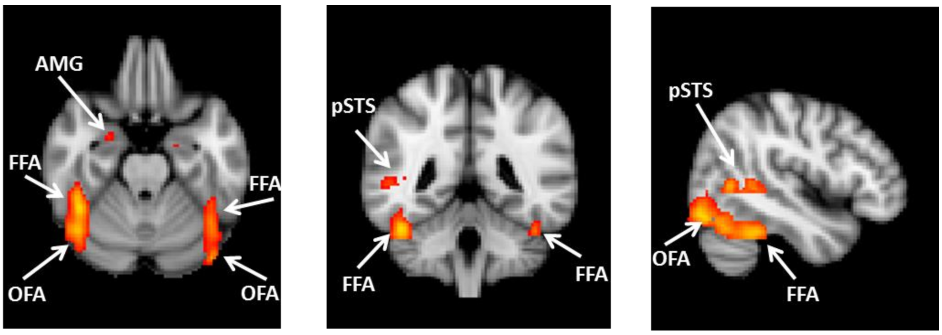


Figure 5.3 Location of face selective regions. Average ROIs across all participants transformed into standard space.

Table 5.1 MNI (mm) coordinates and thresholds of group-level face responsive (OFA, FFA) clusters

Region	Hemisphere	x	y	z	Threshold (Z)
FFA	L	-42	-56	-26	2.86
	R	44	-48	-26	3.99
OFA	L	-42	-80	-20	2.91
	R	46	-76	-16	3.02

5.3.1 Analyses with individual-defined ROIs

A mixed-ANOVA with Participant Group (Chinese, Western Caucasian) as the between-subjects factor, and Face Ethnicity (Chinese, Western Caucasisan), Condition (sIsE, sIdE, dIsE, dIdE), and Hemisphere (left, right) as within-subjects factors was conducted for each ROI (FFA, OFA), separately. In the OFA, neither the main effect of Hemisphere ($F(1,17) = 0.8$, $p > .1$, partial $\eta^2 = 0.05$), nor the interactions of Hemisphere with Condition, Face Ethnicity or Participant Group (Hemisphere x Condition: $F(3,51) = .02$, $p > .1$, partial $\eta^2 = 0.001$; Hemisphere x Condition x Participant Group: $F(3,51) = 2.3$, $p = .09$, partial $\eta^2 = 0.12$; Hemisphere x Condition x Face Ethnicity: $F(3,51) = 1.7$, $p > .1$, partial $\eta^2 = 0.1$; Hemisphere x Condition x Participant Group x Face Ethnicity: $F(3,51) = 2.4$, $p = .08$, partial $\eta^2 = 0.12$) were significant. In the FFA, only the main effect of Hemisphere was significant, $F(1,19) = 11.4$, $p < .01$, partial $\eta^2 = 0.38$, showing an right hemisphere advantage. None interaction of Hemisphere with other factors were significant. Therefore, in the following analyses, left and right hemispheres were combined across hemispheres for all ROIs.

A mixed-ANOVA with Participant Group (Chinese, Western Caucasian) as the between-subjects factor, and Face Ethnicity (Chinese, Caucasisan), Condition (sIsE, sIdE, dIsE, dIdE) as within-subject factors was conducted for each ROI (FFA, OFA), separately. Figure 5.4 shows peak responses to different conditions in each ROI. For the FFA, there was a significant main effect of Condition, $F(3,69) = 5.8$, $p < .001$, partial $\eta^2 = 0.21$. This was due to higher responses in the different identity and different expression (dIdE) condition compared to the other three conditions (dIdE vs. sIsE: $t_{(24)} = 4.3$, $p < .01$; dIdE vs. sIdE: $t_{(24)} = 3.4$, $p < .05$; dIdE vs. dIsE: $t_{(24)} = 3.6$, $p < .01$). The main effect of Condition was also modulated by a two-way interaction of Condition x Face Ethnicity, $F(3,69) = 3.1$, $p < .05$, partial $\eta^2 = 0.12$. This reflected a greater difference between conditions for the Chinese faces compared to the Western Caucasian faces. However, there was no main effect of Group ($F(1,23) = 0.001$, $p > .1$, partial $\eta^2 = 0.0$) or any interactions with Group.

In the OFA, there was a significant main effect of Condition, $F(3,66) = 3.8$, $p < .05$, partial $\eta^2 = 0.15$. This was due to a higher response in the dIdE condition compared to the sIsE condition, $t_{(23)} = 3.3$, $p < .05$. There was no main effect of Group ($F(1,22) = 2.9$, $p > .1$, partial $\eta^2 = 0.12$) or any significant interactions.

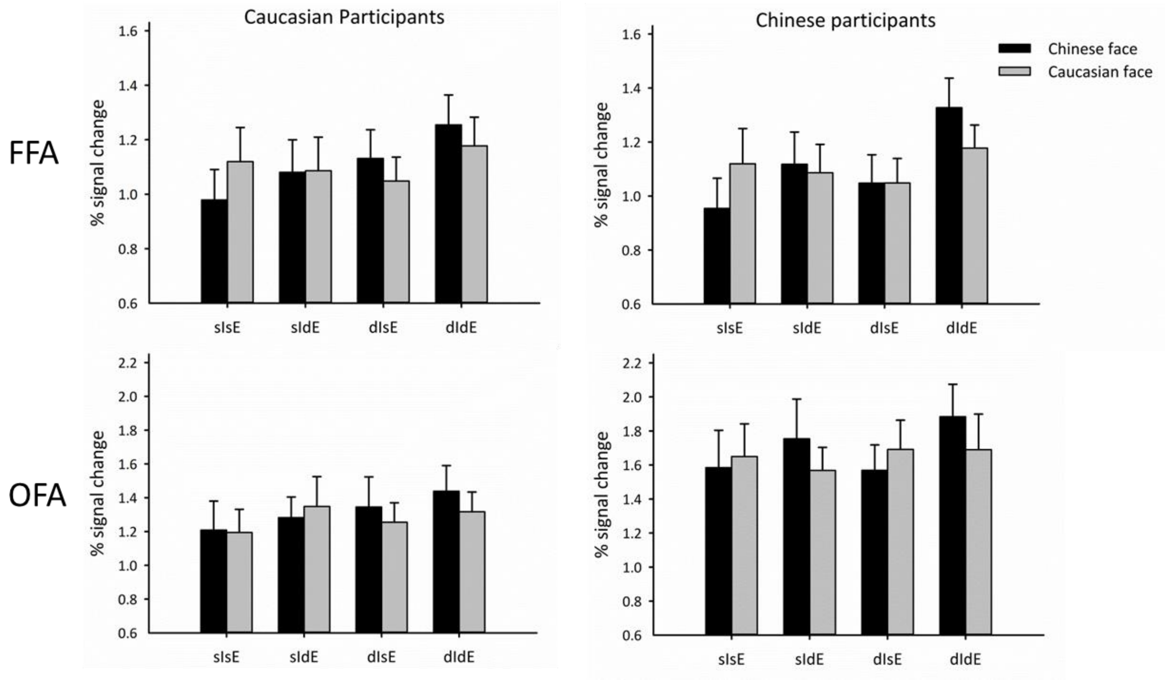


Figure 5.4 Peak responses to different conditions in the FFA and OFA defined by individuals (with standard error bars).

5.3.2 Analyses with group-level ROIs

The same mixed-ANOVAs (Group x Face Ethnicity x Condition) were conducted for each ROI using the group-level ROIs from the localiser scan. In the FFA, the results showed that the main effect of Condition was not significant, $F(3,78) = 1.3$, $p > .1$, partial $\eta^2 = 0.05$, but there was a significant interaction of Face Ethnicity and Condition, $F(3,78) = 4.7$, $p < .01$, partial $\eta^2 = 0.15$, reflecting a larger adaptation effect with East Asian faces, $F(3,78) = 4.7$, $p < .01$. The analyses found bigger responses in the dIdE condition versus the sIsE condition, $t_{(27)} = 3.8$, $p < .05$. There was also a trend showing stronger activations in the dIsE condition than the sIsE condition, $t_{(27)} = 2.8$, $p = .07$. There was no main effect of Group ($F(1,26) = 1.4$, $p > .1$, partial $\eta^2 = 0.05$) or any significant interactions.

In the OFA, the main effect of the Condition was not significant, either, $F(3,78) = 1.3$, $p > .1$, partial $\eta^2 = 0.05$, but there was a significant interaction of Face Ethnicity and Condition, $F(3,78) = 3.5$, $p < .05$, partial $\eta^2 = 0.12$, indicating also a reliable adaptation effect with East Asian faces. Similarly, responses in the dIdE ($t_{(27)} = 3.3$, $p < .05$) and dIsE ($t_{(27)} = 2.9$, $p < .05$) condition were bigger than the sIsE condition, whilst there was no difference between condition dIdE and dIsE ($t_{(27)} = 0.53$, $p > .1$). There was no main effect of Group ($F(1,26) = 3.0$, $p = .09$, partial $\eta^2 = 0.10$) or any significant interactions.

5.3.3 Whole brain analysis

To find regions outside the core face-selective regions showing sensitivity to facial identity and expression, I performed a whole brain analysis. Consist with the ROI analysis, only face-selective regions were detected. This may reflect a lack of power due to the small number of participants.

5.4 Discussion

The aim of this chapter was to investigate the neural representation of facial identity and expression in own-race and other-race faces. The results showed that core face-selective regions from Haxby et al.'s (2000) model (FFA and OFA) were sensitive to changes of faces in identity and expression, as indicated by the significant decrease in the responses to the same identity and same expression (sIsE) condition compared to the conditions with different identity and different expression (dIdE). These results were consistent with previous studies (Harris et al., 2012; 2014b). However, there were no significant differences in the responses to own-race and other-race faces.

Our results differ from previous studies that have shown differences in the magnitude of the univariate response to own-race and other-race faces in face-selective regions (Golby et al., 2001; Kim et al., 2006). There are a number of possible reasons for why our results are different from previous studies: One possibility is a difference in the task design. Previous studies that have found other-race effects in the FFA have either used recognition (Kim et al. 2006) or memory tasks (Golby et al., 2001). In contrast, the present study used an orthogonal red spot detection task. However, Natu et al. (2011) also used an orthogonal detection task, and found differences in the neural responses of the FFA to own- versus other-race faces. Nevertheless, the results of Natu et al.'s (2011) study are difficult to interpret as the FFA responded more strongly to own- than other-race faces only at the first few time points of the block, but the activation pattern reversed at later time points.

The majority of these findings have been conducted with African-American and Western Caucasian faces. In contrast, this study has used East Asian and Western Caucasian faces. It is possible that attitudes to the faces may play a role in the patterns of response. Support for this possibility comes from studies showing that activation in the FFA can be modulated by higher-level social attitudes to different races of faces. For example, Brosch et al. (2013) showed that the successful prediction of responses to own- and other-race faces on the basis of BOLD activation patterns in the FFA was restricted to individuals with high pro-

White bias. Van Bavel and colleagues (2008) also found that the other race effect in fusiform activity could stem from the motivated aspects of own race categorization.

Even though no differences the responses to own-race and other-race faces were found in the FFA and OFA, it could be that the univariate analysis was not sensitive enough to determine the reliability of the response patterns between different face conditions (Haxby et al., 2001). However, as this study was designed for a block adaptation task, it was not appropriate to conduct MVPA. With multi-voxel analyses, Brosch and colleagues (2013) have found that a pattern classifier applied to voxels in the FFA was able to discriminate the brain activity map to own- versus other-race faces. Although Natu and colleagues (2011) were unable to find an own-race bias in the pattern of response in the FFA, an own-race bias was evident when a broader range of ventral temporal areas were involved. Their results indicated that regions outside of traditional face-selective areas could also provide useful information to determine the race of a face.

The present study aimed also to investigate the neural representation of faces of different races in multiple face-selective regions, such as the OFA, the STS, and the amygdala. However, the activations to stimuli were quite different among participants, and regions like the STS and the amygdala were only detected in a small number of the participants. This fact therefore restricted us mainly to focus on the brain responses in the FFA. In order to achieve a systematic viewpoint regarding the brain representations to difference races of faces in these other face-selective regions, a larger sample size is needed in further studies.

In conclusion, this chapter found that the FFA and the OFA were sensitive to changes in facial identity and expression, but the impact of the other-race effect in identity and expression processing could not be identified from univariate analyses of fMRI in these regions.

6. General Discussion

Psychologists and philosophers have long suggested that basic processes of cognition and perception such as attention, memory, categorization, and critical analysis are universal across cultures. However, recent studies have found that people from difference cultures, especially East Asians and Western Caucasians, perceive and think about the world in very different ways (Nisbett & Masuda, 2003). One of the main reasons is the different social practice of the two societies. For example, East Asians live in complex social networks with other individuals, with the family and the society. The socially interdependent relations make them always have to coordinate their behaviors with others to maintain social harmony. Instead, the western societies value individualism and autonomy. This leads to individuals from western cultures paying more attention to their personal goals.

The main purpose of this thesis was to investigate the influence of cultures on people's perception of faces and also their underlying mechanisms. To be more specific, the core work attempted to answer the following four questions:

- Are there cross-cultural differences in the perception and recognition of facial expressions of basic emotions?
- What are the potential factors to explain the other-race effect in facial expression recognition?
- What are the extent of cross-cultural similarities and differences in the processing of facial identity and expression of own-race and other-race faces?

- What are the neural mechanisms of the other-race effects in facial identity and expression processing?

To address these questions, fully balanced crossover designs were always used by testing Western Caucasian and East Asian participants with Western Caucasian and East Asian face stimuli. Therefore, the other-race effect was captured by the interaction of participant group and face ethnicity in the ANOVA analysis. This thesis gives us a better understanding that how faces of different races, which are not specific to white and black, are processed cross-cultures.

6.1 Are there cross-cultural differences in the processing of facial expressions of basic emotions?

Facial expressions of several basic emotions have long been considered to be processed universally, as suggested to serve evolved biological functions pertinent to survival (Darwin, 1872; Ekman, 1980). For example, facial expression of fear has been suggested to indicate potential physical threat in the environment. Studies have found that even people living in preliterate cultures who have minimal exposure to literate cultures pose facial expressions in a similar way to expressions posed by people from the western cultures (Ekman, 1972). Although there are differences in the recognition rates of expressions by members from different cultures, the performance is always above chance levels.

Even though people are able to recognize facial expressions across cultures, researchers have argued that the cultural differences cannot be ignored; people show better recognition of the expressions posed by their own group members (Elfenbein & Ambady, 2002b). Recent studies have aimed to find the way facial expressions of different races are processed by people from different cultural backgrounds, using eye-tracking and reverse correlation methods (Jack et al., 2009; 2012). These results showed that East Asians mainly focus on the eye regions when recognizing facial expressions, while Western Caucasians tend to use the information from both the eye and mouth regions.

The first study of this thesis therefore systematically investigated the potential cultural similarities and differences in the processing of facial expressions of five basic emotions (anger, disgust, fear, happiness, and sadness). Most of the studies mentioned above only use forced-choice tasks by asking participants to identify the expressions they see, but results with recognition performance only cannot explain the whole story of expression processing. No studies have ever clarified whether or not the divergence starts at the perception of expressions themselves, or it happens later when recognition is involved. To be specific, it is not clear whether culture can shape the perception of expressions, even though previous studies have showed that culture shapes many ways people thinking and processing the world, including categorization, causal attribution, reliance on rules, use of logic, attention, and perception (Nisbett & Masuda, 2003).

Therefore, in a series of experiments, two behavioral tasks were used to examine the perception and recognition of facial expressions posed by Western Caucasian and Chinese actors in Western Caucasian and Chinese participants. The results from Chapter 2 revealed a difference in the influence of cultures on the two types of expression processing. The perceptual similarity judgements of facial expressions across the two groups of participants showed substantial consistencies with correlation coefficients above 0.9. This indicates that people from different cultures see facial expressions in much the same way. In contrast, there were reliable differences (5%-9%) in judgements of expression recognition when participants were required to identify the same facial expressions as used in the perception task. The other race effect for recognizing expression is consistent with previous studies (Ekman, 1972; Jack et al., 2009). However, it is important to note that compared to the magnitude of cross-cultural similarities in perception of expressions, the differences in expression recognition were quite small.

Chapter 2 also investigated whether Western Caucasian and Chinese people use different facial signals to process expressions by only showing either the upper or lower half of the faces. The results showed that perception of expressions from either half of the faces were still consistent across cultures, further confirming that there is no difference in the perception of facial expressions, even to the perception of local features only (e.g., eyes, mouth). The results in the categorization task however showed that the own-group advantage in expression recognition was driven by the fact that people could make better use of information from the mouth regions of their own-race members, which was inconsistent with previous studies suggesting that Chinese people only use information from the eye region to recognize facial expressions (Jack et al., 2012).

Therefore, the answer to the question about whether or not there are cultural differences in processing facial expressions of basic emotions is "no", when considering only the perception of the expressions, but "yes" for the recognition of expressions. The results indicated that the own-group advantages possibly reflected minor cultural 'stylistic' differences in the way in which these emotions are expressed around a common overall template. The organization of the facial muscles makes the lower part of the face relatively mobile compared to the more limited range of movements possible in the eye region, hence more capable of developing such differences. However, compared to the magnitude of cross-cultural agreement, the own-group categorization advantage is quite small.

6.2 What are the potential factors that could explain the own-group advantages in expression processing?

The other-race effect in facial identity processing has been widely found in many previous studies, and there are mainly two theories to explain the cross-cultural differences. The advantages in processing own-race faces could be driven by different levels of experiences with difference races of faces. The perceptual learning theories address the increased contact with own-group members and thus more efficiency at discriminating among own-race members (Furl et al., 2002; Meissner & Brigham, 2001). The social cognitive theories emphasize that the own-group advantages are caused by higher motivation to process own-race faces versus other-race faces (Hugenberg et al., 2010; Sporer, 2001). The own-race faces are processed at a more detailed individuation level, whereas the other-race faces only involves a quick and automatic categorization processing.

This thesis did not directly examine whether or not the two above-mentioned theories could explain the own-group advantages in expression perception. The findings of Chapter 2 that the own-group advantages in facial expression identification were largely restricted to the lower part of the face however make the social motivation explanation an unlikely candidate here. The fact that no differences in reaction times between own-race and other-race face categorization were detected also restricted the possibility of social motivation theories. However, this is not to deny the possible important role of differences in social motivations to own-race and other-race members in perception of faces, as studies have shown that own-group advantages could be created by modulating participants' motivation in the experiment (Young & Hugenberg, 2000). In addition, it is possible that race issues in North America are more sensitive compared to the issues in European countries. Studies have shown that when required to process White and Black faces, White participants in America tend to show more negative bias towards Black than White faces (Brosch et al., 2013; Phelps et al., 2000). The social cognition theories may therefore more fit with studies investigating cultural differences between White and Black faces in cultures where differences in social attitudes to own-race and other-race faces are more pronounced.

Following the findings of Chapter 2, Chapter 3 of this thesis investigated another two possible reasons that could explain the own-group advantages: (1) confusability between certain emotions; (2) differences in the engagement of holistic processing. The results only found own-group advantages in recognition of anger and disgust (and not fear and surprise), even though fear and surprise were found to be more easily confused with each other than anger and disgust. Chapter 3 also found that the differences in emotion recognition could not be explained by differences in holistic processing; both Western

Caucasian and Chinese participants showed a similar amount of expression-composite effect to both races of faces.

Based on the findings of Chapter 2 and Chapter 3, the own-group advantages in face perception could reflect relatively minor cultural differences in the way certain emotions are expressed and interpreted (see also Ekman, 1972; Matsumoto, 1989). Differences in the perceptual experiences to own-race versus other-race faces would then explain the own-group advantages. Recent studies offer some evidence showing that the meaning of disgust can be expressed differently in different cultures. The English word disgust refers to more mixed reactions to both physical and moral disgust scenarios, compared to the Korean and Malayalam words (Han et al., 2015; Yoder et al., 2016).

Nisbett and Masuda (2003) have proposed that culture affect many ways people think and process the world, including perception, categorization, attention, causal attribution, reliance on rules and use of logic. Recent studies also found some neural evidence of cultural influences on people's processing of the world (Goh et al., 2010; Han & Ma, 2014). These cultural differences are thought to be largely driven by people living in different cultural backgrounds with different social practices. For example, East Asians are involved with multiple and complex social relations with other individuals, including family members and society members. They then perceive themselves as embedded within a larger group, and should perform consistently with other individuals within the same group. In contrast, Western Caucasians are relatively independent, involved with less complex social relations than East Asians. They value individualism and autonomy, and are therefore more likely to attend to objects according to their personal goals. The own-group advantage in expression perception could also be explained by this theory. Compared to Western Caucasians, East Asians who live in a collectivistic culture may have learned to suppress negative emotions and therefore show differences in posing and explaining these emotions that could harm social harmony and interpersonal relationships (Matsumoto, 1989; Matsumoto & Ekman, 1989). These kinds of differences in expression and explanation of emotions lead them to be more accurate at recognizing expressions posed by their own-race members than expressions of other-race members.

6.3 "Do they (other-race faces) look all the same to me?"

"Other things being equal, individuals of a given race are distinguishable from each other in proportion to our familiarity, to our contact with the race as whole. Thus, to the uninitiated American all Asiatics look alike, while to the Asiatics, all White men look alike (Feingold, 1914)."

The fact that the other-race effect has been found in a lot of studies has contributed to the impression that there are large cross-cultural differences in face recognition, as reflected in the sentences quoted above from Feingold in the early 90s. One recent study also showed that when people were presented with pairs of faces of either the same or different identities while ERPs were recorded, only the same-race faces elicited the adaptation effect when the two faces presented were the same versus different, indicating that the brain can only tell the differences between the faces of own race members (Vizioli et al., 2010).

In addition, a number of studies have shown other-race effects in the perception of facial expression and identity (Meissner & Brigham, 2001; Elfenbein & Ambady, 2002b; Jack et al., 2012), but no studies have investigated cross-cultural processing of facial identities and expressions simultaneously with structure matched tasks. Hence it remains unclear whether this reflects a common bias in face recognition or independent processes. In Chapter 4, I therefore used a sorting task (Jenkins et al., 2011) to measure the ORE for these two domains. By mapping participants' sorting solutions with response matrices that capture all information of the data and then comparing between the two, I was able to investigate the magnitude of cross-cultural differences and similarities in facial identity and expression processing. The range of the correlation values from 0.7 to 0.91 indicated that even though there was a reliable other-race effect for both groups of participants, their sorting solutions showed high consistency across cultures. These finding means that the widely accepted idea that the other-race effect is large is clearly overstated. And this was true for both the facial identity and expression processing, as the results found very similar results across these two domains.

6.4 The neural representations of cross-cultural differences in the human brain

Having found reliable cross-cultural differences in facial identity and expression processing, Chapter 5 followed up to investigate the underlying neural mechanisms of the ORE. Previous studies have explored the potential neural correlates in the Haxby et al.'s (2000) core face-selective regions with both the univariate and multivariate analyses, but the results were inconsistent. One possible reason is that different tasks were used among studies. The strong activation to own-race compared to other-race faces in the FFA have been found in studies with recognition (Kim et al., 2006), memory (Golby et al., 2001), or orthogonal detection task (Natu et al., 2011). Another possible reason is that the involvement of social attitude to different races of faces (especially to the white and black faces) could also modulate the differences in the FFA activations found in previous studies (Brosch et al., 2013; Van Bavel et al., 2008).

Chapter 5 of this thesis used one orthogonal detection task, the red-dot task, to examine the neural activities to different race of faces when the stimuli were only passively viewed. And by using the powerful fMR-adaptation technique (Grill-Spector et al., 2006), this study could tell the sensitivity of the neural representation to certain changes in the stimulus. At first, a ROI analysis found that both the FFA and the OFA were sensitive to the changes in facial identity and expression of both the Western Caucasian and East Asian faces. There were reliable attenuations in the neural responses to the repetition of faces (adaptation effect). These results were consistent with previous studies showing that the FFA and the OFA could tell the changes in facial identity and expression (Harris et al., 2012; 2014b).

However, the results did not find the differences in the adaptation effects to own- versus other-race faces, which was inconsistent with the hypothesis. According to the perceptual learning theories, the long-term differences in the perceptual experiences to own-race compared to other-race faces should lead to a bigger adaptation to own-race faces. The results therefore indicated that just passively processing the faces, without the involvement of further recognition or memory processing, would not lead to significant differences in the activations to faces of different races in these core face regions. Differences in activations to own-race and other-race faces in the FFA, OFA and STS were also found with different attentional task (identification or categorization task) (Liu et al., 2014). However, it could also be that the ROI analysis was not sensitive enough to determine the reliability of the response patterns between different conditions (Haxby et al., 2001), and this could be the next step of further studies.

Previous studies have suggested mainly two theories to explain the other-race effect: the perceptual learning theories (Golby et al., 2001; Kim et al., 2006) and the social cognitive theories (Brosch et al., 2013; Van Bavel et al., 2008). However, these theories are not conflict with each other, as they mainly focus on different phases of face processing: the former emphasizes the differences in perceptual expertise to different races of faces that lead to the differences in the neural representations of faces, while the latter stresses the impact of evaluations of faces related to social categorization on face processing. Many previous studies have showed that brain regions such as the amygdala are sensitive to people's attitude to own-race compared to other-race members. There would be greater BOLD responses to other-race faces when people have greater racial bias to other-race members (Phelps et al., 2000; Wheeler & Fiske, 2005). By examining more related regions, such as the amygdala, the core face regions the FFA, the OFA, and the STS, and also their functional connections among each other, future studies would achieve a better understanding of the neural correlates of the other-race effect.

6.5 What are the relationships between the processing of facial identity and expression?

The question of whether or not the processing of facial identity and expression is independent has triggered a lot debate. Many neural studies have shown that the invariant (e.g., facial identity) and changeable (e.g., facial expression) aspects of faces were represented in different regions in human brain (Haxby et al., 2000; Fox et al., 2009; Harris et al., 2014b). However, this does not indicate that the processing of these two domains is independent with each other. One good way of investigating this question was to present these two aspects of information simultaneously and observe whether or not the processing of one aspect is affected by another. Behavioural studies have shown that processing of these two domains is indeed influenced with each other, but these interactions are not equivalent. The processing of facial expressions was affected by the changes in irrelevant facial identity, while changes in expression instead had no influence on the recognition of facial identity (Schweinberger & Soukup, 1998).

By investigating the cross-cultural processing of facial identity and expression at the same time, this thesis can also give us a hint about their relationship from the cultural processing aspect. Using a novel behavioral task, Chapter 4 of this thesis found that there were reliable own-group advantages at processing own- versus other-race facial identity and expression, which indicated that the processing of own- versus other-race faces across these two domains shares some similarities. However, the factors could lie in either the early perceptual or later processing, which could be affected by social factors.

Chapter 2 first investigated two different ways of processing facial expressions posed by different race actors. The results however only found the own-group advantages at the recognition of expressions, rather than perception. This was inconsistent with the findings in identity processing, which showed that the preferences to own-race faces started even early in the perception processing (Lindsay et al., 1991; Megreya et al., 2011; Walker & Tanaka, 2003), not even to say the later recognition processing.

Further studies in Chapter 3 also indicated some differences in the processing of these two domains. The results showed that the own-group advantages in expression processing could not be explained by the differences in the holistic processing of faces, which was instead a widely proved phenomenon in identity processing. On the contrary, the cross-cultural differences in expression processing more likely exist in the later recognition phase that found to be affected by social factors, such as specific cultural rules in the decoding of facial expressions.

Therefore, the results of this thesis show us some differences in the cross-cultural processing of facial identity and expression. The preference to own-race faces happens very early at the perceptual level in identity processing, while relatively late at the recognition level in expression processing. In sum, cross-cultural studies can offer an interesting perspective on the relationship in the processing of facial identity and expression. From the cultural processing aspect, the way culture shapes the processing of facial identity and expression is different and also starts at different stages. To give a comprehensive exploration of this question, further studies should start with investigating the underlying processing mechanisms of these two aspects in the human brain, and should maybe also link back to the ecological meanings of these two different face signals to the survival of us human beings.

6.6 Limitations of present thesis

One limitation of this present thesis is that facial stimuli from two face sets (the KDEF set and the Chinese face set) are not fully matched in quality, as the two sets of images were not created from the same lab. Even though these images were created by asking participants to pose strong and clear expressions, the overall recognition rates for Western Caucasian faces were relatively higher than Chinese faces (even though the physical properties, such as the image size and lamination values, have been adjusted and balanced by Photoshop CS3 software before testing). These differences in images have made the ANOVA results a bit more complicated, showing more interactions of face ethnicity with other factors that were not the main focus of interest and also making some caveats to the explanation of the results. However, the statistical analyses had found significant interaction of face ethnicity by participant group, indicating the own-group advantages in face perception. Future researches should use images in good qualities or at least quality-matched to explore the cross-cultural processing of faces.

In addition, the small sample size of the imaging study has limited the explanation of the results. Owing to time restrictions, only fourteen participants were recruited in each group. In the ROI analysis, regions like the rSTS and amygdala were only detected in a small number of participants, and will not therefore give a reliable result in further analyses. This restricted the analyses only to focus on the responses in the OFA and the FFA, which are mainly sensitive to facial identity. With a larger sample size, future studies can have a better understanding of the neural representation of own-race and other-race faces in the expression-sensitive regions, such as the STS and the amygdala. The task design used in the fMRI study also restricted the explanation of the results. Future researches can use a different task design, for example, the event-related design, to further explore the neural mechanisms of the other-race effect.

6.7 Future directions

The work of this thesis mainly explores similarities and differences in cross-cultural perception of face expressions. The results indicate that the own-group advantage in expression perception possibly reflect minor cultural 'stylistic' differences in the way certain emotions are expressed around a common overall template, and also in the way people from different cultures explain them. Future research should attempt to quantify these differences in the expression and explanation of emotions. One way to explore how these expressions are posed would be by morphing the expression images of each category together according to participants' subjective responses to measure the differences directly.

By systematically manipulating the texture-based and shape-based cues available in face stimuli, future studies could also examine how different facial cues are used by people from different cultures to perceive facial expressions posed by their own-race and other-race members. Shape information indicates the spatial locations corresponding to facial features, and also the shape of facial features themselves, whereas surface information indicates pattern of surface pigmentation resulting from local changes in the reflectance properties of the skin (Bruce & Young, 1998, 2012). Even though shape-based and texture-based information have been found to contribute differently to the perception of facial identity and expression (Bruce & Young, 1998; Harris et al., 2014b; Calder, Young, Perett, Etcoff, & Rowland, 1996), studies also showed that optimal performance is achieved when both types of information are available (Calder, Burton, Miller, Young, & Akamatsu, 2001; Michel et al., 2013). Michel and colleagues (2013) examined the contribution of these two types of cues to the cross-cultural processing of facial identity. Their results showed that East Asian and Western Caucasian participants relied differently on shape and texture information to recognize faces (both own-race and other-race faces). By creating shape varying only (consistent texture) and texture varying only (consistent shape) images, future studies could also investigate how information is used by people from different cultures when recognizing facial expressions, and how this is represented in the brain.

6.8 Conclusions

This thesis aimed to provide a better understanding of the influences of culture on people's perception of faces, especially the facial identity and expression. Similar to the well-known phenomenon of the other-race effect in identity processing, this thesis also found a reliable own-group advantage in facial expression processing. However, contrary to the fact that the other-race effect is found in both the perception and recognition of facial identity, only difference in the recognition of facial expressions of basic emotions was found. Further studies found that this cross-cultural difference in expression processing was only driven by differences in information using from the lower part of faces; both the Western Caucasian and Chinese participants could make better use of the information from mouth regions posed by their own race members. The subsequent study found that the difference in expression recognition was only found in anger and disgust but not surprise and fear, and there was no difference in the holistic processing to own- versus other-race expressions. In addition, even though the cross-cultural differences in identity and expression processing have been found to be a reliable phenomenon, it is very small and there is actually a very large cross-cultural agreement, indicating the widely accepted idea that the cross-cultural difference is huge (“they all look the same”) is overestimated. Finally, the fMRI study further found that the impact of the other-race effect in identity and expression processing could not be identified from univariate analyses of fMRI in the core face-selective regions in Haxby et al’s (2000) model.

Appendices

Appendix1

Please write down the Chinese meaning of these five expressive words:

请写出下面情绪单词的中文意思：

Anger __________________________

Disgust __________________________

Fear __________________________

Happy __________________________

Sad __________________________

Appendix 2

Questionnaire for study of face perception by Chinese participants

Gender: ___________________

Age (years): ___________________

1. Which part of China are you from?

   __________ (Province, or Special Administrative Region)

2. How long have you lived in China?

   __________ (Years)

3. How long have you been in the UK?

   __________ (Years) __________ (Months)

Thank you for helping with this study. If you have any questions about the purpose of the study or this questionnaire, please contact Xiaoqian Yan (xy760@york.ac.uk).

中国被试面孔感知研究调查问卷

性别：　　　　____________________

年龄（岁）：　____________________

1. 你来自中国的哪个地区？

   _____________（省，或者特别行政区）

2. 你在中国住了多久？

   _____________（年）

3. 你来英国多久了？

   ______________（年）______________（月）

非常感谢参加本次实验，如果你对本问卷调查的目的存在任何疑问，请与闫晓倩联系（xy760@york.ac.uk）。

Appendix 3

Table 1 Average recognition accuracies (with standard errors) for each expression (Chinese and KDEF faces) in Chinese and Western Caucasian participants in Chapter 2 (Experiment 3) and Chapter 3 (Experiment 1)

		Anger		Disgust		Fear		Happiness		Sadness		Surprise	
		Chinese Faces	KDEF Faces	Chinese Faces	KDEF Faces	Chinese Faces	KDEF Faces	Chinese Faces	KDEF Faces	Chinese Faces	KDEF Faces	Chinese Faces	KDEF Faces
Chapter 2	Chinese Participants	0.71 (0.19)	0.67 (0.17)	0.67 (0.24)	0.57 (0.18)	0.82 (0.17)	0.70 (0.15)	0.96 (0.06)	0.96 (0.05)	0.69 (0.20)	0.85 (0.10)	-	-
	Western Caucasian Participants	0.63 (0.11)	0.83 (0.17)	0.57 (0.16)	0.64 (0.15)	0.85 (0.16)	0.69 (0.19)	0.95 (0.04)	0.95 (0.06)	0.58 (0.17)	0.86 (0.11)	-	-
Chapter 3	Chinese Participants	0.76 (0.04)	0.65 (0.04)	0.73 (0.04)	0.62 (0.04)	0.49 (0.04)	0.46 (0.04)	0.93 (0.02)	0.95 (0.01)	0.62 (0.05)	0.84 (0.02)	0.79 (0.02)	0.79 (0.03)
	Western Caucasian Participants	0.65 (0.03)	0.84 (0.03)	0.61 (0.04)	0.63 (0.03)	0.52 (0.04)	0.46 (0.03)	0.93 (0.01)	0.95 (0.01)	0.61 (0.04)	0.88 (0.02)	0.80 (0.03)	0.85 (0.02)

References

Adolphs, R., Tranel, D., Hamann, S., Young, A. W., Calder, A. J., Phelps, E. A.,…, & Damasio, A. R. (1999). Recognition of facial emotion in nine individuals with bilateral amygdala damage. *Neuropsychologia, 37*, 1111−1117.

Andrews, T. J., Davies-Thompson, J., Kingstone, A., & Young, A. W. (2010). Internal and external features of the face are represented holistically in face-selective regions of visual cortex. *The Journal of Neuroscience, 30*(9), 3544–3552.

Andrews, T. J., & Ewbank, M. P. (2004). Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *NeuroImage, 23*, 905−913.

Arsalidou, M., Morris, D., & Taylor, M. J. (2011). Converging evidence for the advantage of dynamic facial expressions. *Brain Topography, 24*, 149−163.

Baudouin, J. Y., & Humphreys, G. W. (2006). Configural information in gender categorisation. *Perception, 35*(4), 531−540.

Beaupré, M. G., & Hess, U. (2005). Cross-cultural emotion recognition among Canadian ethnic groups. *Journal of Cross-Cultural Psychology, 36*(3), 355−370.

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The cross-category effect mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science, 18*(8), 706−712.

Biehl, M., Masumoto, D., Ekman, P., Hearn, V., Heider, K., Kudoh, T., & Ton, V. (1997). Matsumoto and Ekman's Japanese and Western Caucasian Facial Expressions of

Emotion (JACFEE): reliability data and cross-national differences. *Journal of Nonverbal Behavior, 21*(1), 3−21.

Blair, R. J. R., Morris, J. S., Frith, C. D., Perrett, D. I., & Dolan, R. J. (1999). Dissociable neural responses to facial expressions of sadness and anger. *Brain, 122*(5), 883−893.

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS one, 3*(8), e3022.

Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *The Journal of Neuroscience, 16*(11), 3737−3744.

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

Brigham, J. C., Bennett, L. B., Meissner, C. A., Mitchell, T. L. (2007). The influence of race on eyewitness memory. In: Lindsay, R.C.L., Ross, D.F., Read, J.D., Toglia, M.P. (Eds), *The Handbook of Eyewitness Psychology, Volume II: Memory for People* (pp. 257–281). Mahwah, NJ: Lawrence Erlbaum Associates.

Brosch, T., Bar-David, E., & Phepls, E. A. (2013). Implicit race bias decreases the similarity of neural representations of black and white faces. *Psychological Science, 24*(2), 160-166.

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J. B., Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied, 5*, 339–360.

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

Bruce, V., & Young, A. (1998). *In the eye of the beholder: The science of face perception.* Oxford: Oxford University Press.

Bruce, V., & Young, A. W. (2012). *Face perception*. Psychology Press.

Bruyer, R., & Crispeels, G. (1992). Expertise in person recognition. *Bulletin of the Psychonomic Society, 30(6), 501*−504.

Calder, A. J., Burton, A. M., Miller, P., Young, A. W., & Akamatsu, S. (2001). A principal component analysis of facial expressions. *Vision Research, 41*(9), 1179−1208.

Calder, A. J., & Jansen, J. (2005). Configural coding of facial expressions: The impact of inversion and photographic negative. *Visual Cognition, 12*(3), 495−518.

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience, 6*, 641−651.

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*(2), 527–551.

Calder, A. J., Young, A. W., Perrett, D. I., Etcoff, N. L., & Rowland, D. (1996). Categorical perceptions of morphed facial expressions. *Visual Cognition, 3*(2), 81−117.

Calvert, K. L., Doar, M. B., & Zegura, E. W. (1997). Modeling internet topology. *Communications Magazine, IEEE, 35*(6), 160−163.

Calvo, M. G., & Lundqvist, D. (2008). Facial expressions of emotion (KDEF): Identification under different display-duration conditions. *Behavior Research Methods, 40*, 1, 109−115.

Campbell, J., & Burke, D. (2009). Evidence that identity-dependent and identity-dependent neural populations are recruited in the perception of five basic emotional facial expressions. *Vision Research, 49*, 1532−1540.

Chance, J., & Goldstein, A. (1996). The other-race effect and eyewitness identification. In Sporer, S. L., Malpass, R. & Koehnken, G. (Eds.), *Psychological issues in eyewitness identification* (pp. 153–176). Mahwah, NJ: Lawrence Erlbaum Associates.

Chiao, J. Y., Iidaka, T., Gordon, H. L., Nogawa, J., Bar, M., Aminoff, E., ..., & Ambady, N. (2008). Cultural specificity in amygdala response to fear faces. *Journal of Cognitive Neuroscience, 20*(12), 2167−2174.

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

Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of black and white faces. *Psychological Science, 15*(12), 806−813.

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in cognitive sciences, 7*(9), 415−423.

Darwin, C. (1904). *The expression of the emotions in man and animals*. London, UK: Murray. (Original work published 1872)

Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of expertise. *Journal of Experimental Psychology: General, 115*(2), 107−117.

Du, S. C., Tao, Y., & Martinex, A. M. (2014). Compound facial expressions of emotion. *Proceedings of the National Academy of Sciences, 111*(15), E1454-E1462.

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska Symposium on Motivation,* 1971 (Vol.19, pp. 207−282). Lincoln: University of Nebraska Press.

Ekman, P. (1980). *The face of man: Expressions of universal emotions in a New Guinea village*. New York: Garland STPM Press.

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin, 115*, 268−287.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect.* Palo Alto, California: Consulting Psychologists Press.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behaviour: Categories, origins, usage, and coding. *Semiotica, 1*, 49−98.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: Investigator's guide.* Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-culture elements in facial display of emotions. *Science, 164* (3875), 86-88.

Elfenbein, H. A., & Ambady, N. (2002a). Is there an in group advantage in emotion recognition? *Psychological Bulletin, 128*, 243−249.

Elfenbein, H. A., & Ambady, N. (2002b). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin, 128*, 203−235.

Elfenbein, H. A., & Ambady, N. (2003a). Cultural similarity's consequences: A distance perspective on cross-cultural differences in emotion recognition. *Journal of Cross-Cultural Psychology, 34*(1), 92−110.

Elfenbein, H. A., & Ambady, N. (2003b). Universals and cultural differences in understanding emtions. *Current Directions in Psychological Science, 12*, 159−164.

Elfenbein, H. A., & Ambady, N. (2003c). When familiarity breeds accuracy: cultural exposure and facial emotion recognition. *Journal of Personality and Social Psychology, 85*(2), 276−290.

Ellamil, M., Susskind, J. M., Anderson, A. K. (2008). Examinations of identity invariance in facial expression adaptation. *Cognitive, Affective, & Behavioral Neuroscience, 8*(3), 273−281.

Engell, A. D., & Haxby, J. V. (2007). Facial expression and gaze-direction in human superior temporal sulcus. *Neuropsychologia, 45*(14), 3234−3241.

Farkas, L. G., Katic, M. J., & Forrest, C. R. (2005). International anthropometric study of facial morphology in various ethnic groups/races. *Journal of Craniofacial* Surgery, 16(4), 615−646.

Feingold, C. A. (1914). The influence of environment on identification of persons and things. *Journal of the American Institute of Criminal Law and Criminology, 5*, 39−51.

Feng, L., Liu, J. G., Wang, Z., Li, J., Li, L., Ge, L. Z., …, & Lee, K. (2011). The other face of the other-race effect: An fMRI investigation of the other-race face categorization advantage. *Neuropsychologica, 49*(13), 3739−3749.

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

Flack, T. R., Andrews, T. J., Hymers, M., Al-Mosaiwi, M., Marsden, S. P., Strachan, J. W. A., …, & Young, A. W. (2015). Responses in the right posterior superior temporal sulcus show a feature-based response to facial expression. *Cortex, 69*, 14−23.

Fox, C. J., & Barton, J. J. S. (2007). What is adapted in face adaptation? The neural representations of expression in the human visual system. *Brain Research, 1127*, 80−89.

Fox, C. J., Oruç, I., & Barton, J. J. (2008). It doesn't matter how you feel. The facial identity aftereffect is invariant to changes in facial expression. *Journal of Vision, 8*(3), 11, 1−13.

Fox, C. J., Moon, S. Y., Iaria, G., & Barton, J. J. (2009). The correlates of subjective perception of identity and expression in the face network: An fMRI adaptation study. *NeuroImage, 15*(44), 569−580.

Freeman, J., & Ziemba, C. M. (2011). Unwrapping the ventral stream. *The Journal of Neuroscience, 31*(7), 2349−2351.

Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception, 29*, 159−170.

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

Ganel, T., & Goshen-Gottstein, Y. (2004). Effects of familiarity on the perceptual integrality of the identity and expression of faces: The parallel-route hypothesis revisited. *Journal of Experimental Psychology: Human Perception and Performance, 30*(3), 583−597.

Ganel, T., Valyear, K. F., Goshen-Gottstein, Y., & Goodale, M. A. (2005). The involvement of the "fusiform face area" in processing facial expression. *Neuropsychologia, 43*(11), 1645−1654.

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training 'greeble' experts: a framework for studying expert object recognition processing. *Vision Research, 38*, 2401−2428.

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanism for expert object recognition: bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance, 28*(2), 431−436.

George, N., Driver, J., & Dolan, R. J. (2001). Seen gaze-direction modulates fusiform activity and its coupling with other brain areas during face processing. *NeuroImage, 13*, 1102–1112.

Glascher, J., Adolphs, R. (2003) Processing of the arousal of subliminal and supraliminal emotional stimuli by the human amygdala. *The Journal of Neuroscience, 23*, 10, 274–10,282.

Goh, J. O. S., Leshikar, E. D., Sutton, B. P., Tan, J. C., Sim, S. K. Y., Hebrank, A. C., & Park, D. C. (2010). Culture differences in neural processing of faces and houses in the ventral visual cortex. *Social Cognitive and Affective Neuroscience, 5*(2-3), 227–235.

Golby, A. J., Gabrieli, J. D., Chiao, J. Y., & Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. *Nature neuroscience, 4*(8), 845–850.

Gong, X., Huang, Y. X., Wang, Y., & Luo, Y. J. (2011). Revision of the Chinese facial affective picture system. *Chinese Mental Health Journal, 1*(25), 40−46.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology, 74*(6), 1464–1480.

Grèzes, J. (1998). Top down effect of strategy on the perception of human biological motion: a PET investigation. *Cognitive Neuropsychology, 15*(6-8), 553–582.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in cognitive sciences, 10*(1), 14−23.

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience, 7*(5), 555−562.

Gur, R. E., et al. (2002). An fMRI study of facial emotion processing in patients with schizophrenia. *American Journal of Psychiatry, 159*, 1992−1999.

Hancock, K. J., & Rhodes, G. (2008). Contact, configural coding and the other-race effect in face recognition. *British Journal of Psychology, 99*, 45−66.

Habel, U., et al. (2007). Amygdala activation and facial expressions: Explicit emotion discrimination versus implicit emotion processing. *Neuropsychologia, 45*, 2369−2377.

Han, D, Kollareth, D., & Russell, J. A. (2015). The words for disgust in English, Korean, and Malayalam question its homogeneity. *Journal of Language and Social Psychology*, 1−20.

Han, S., & Ma, Y. (2014). Cultural differences in human brain activity: A quantitative meta-analysis. *NeuroImage, 99*, 293-300.

Harris, R., Young, A. W., & Andrews, T. J. (2012). Morphing between expressions dissociates continuous from categorical representations of facial expression in the human brain. *Proceedings of the National Academy of Sciences, 109*(51), 21164−21169.

Harris, R. J., Young, A. W., & Andrews, T. J. (2014a). Dynamic stimuli demonstrate a categorical representation of facial expression in the amygdala. *Neuropsychologia*, 56, 47-52.

Harris, R. J., Young, A. W., & Andrews, T. J. (2014b). Brain regions involved in processing facial identity and expression are differentially selective for surface and edge information. *NeuroImage*, *97*, 217-223.

Hart, A. J., Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., & Rauch, S. L. (2000). Differential response in the human amygdala to racial outgroup vs ingroup face stimuli. *Neuroreport, 11*(11), 2351-2354.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science, 293*(5539), 2425–2430.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in cognitive sciences, 4*(6), 223-233.

Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., & Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron, 22*(1), 189-199.

Hay, D. C., & Young, A. W. (1982). The human face. In A. W. Ellis (Ed.), *Normality and Pathology in Cognitive Functions*. London: Academic Press.

Hayward, W. G., Rhodes, G., & Schwaninger, A. (2008). An own-race advantage for components as well as configurations in face recognition. *Cognition, 106*(2), 1017-1027.

Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience, 3*, 80–84.

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychological review, 117*(4), 1168–1187.

Izard, C. E. (1971). *The face of emotion.* New York: Appleton-Century-Crofts.

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*(18), 1543−1548.

Jack, R. E., Caldara, R., & Schyns, P. G. (2012). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology: General, 141*(1), 19−25.

Jack, R. E., Garrod, O. G. B., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences, 109*(19), 7241−7244.

Jenkins, R., White, D., Van Montfort, X., Burton, A. M. (2011). Variability in photos of the same face. *Cognition, 121*(3), 313−323.

Jiang, Y., & He, S. (2006). Cortical responses to invisible faces: Dissociating subsystems for facial-information processing. *Current Biology, 16*, 2023−2029.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40*, 1−19.

Johnson, P., Mayes, A., Hughes, M, & Young, A. W. (2013). Brain networks subserving the evaluation of static and dynamic facial expressions. *Cortex, 49*(9), 2462–2472.

Kanwisher, N. G., McDermott, J. & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*, 4302–4311.

Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Ito, K., …, et al. (1999). The human amygdala plays an important role in gaze monitoring. *Brain, 122*, 779–783.

Kilbride, J. E., & Yarczower, M. (1983). Ethnic bias in the recognition of facial expressions. *Journal of Nonverbal Behavior, 8*(1), 27−41.

Kim, J. S., Yoon, H. W., Kim, B. S., Jeun, S. S., Jung, S. L., & Choe, B. Y. (2006). Racial distinction of the unknown facial identity recognition mechanism by event-related fMRI. *Neuroscience Letters, 397*, 279−284.

Krumhuber, E. G., Kappas, A., & Manstead, A. S. R. (2013). Effects of dynamic aspects of facial expressions: A review. *Emotion Review, 5*(1), 41−46.

Kubota, J. T., Banaji, M. R., & Phelps, E. A. (2012). The neuroscience of race. *Nature neuroscience, 15*(7), 940−948.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Neuroperception: Early visual experience and face processing. *Nature, 410*, 890−890.

Levin, D. T. (2000). Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and cross-race recognition deficit. *Journal of Experimental Psychology: General, 129*, 559−574.

Lindsay, D. S., Jack, P. C. Jr., & Christian, M. A. (1991). Other-race face perception. *Journal of Applied Psychology, 76*, 587−589.

Liu, J. G., Wang, Z., Feng, L., Li, J., Tian, J., & Lee, K. (2014). Neural trade-offs between recognizing and categorization own- and other-race faces. *Cerebral Cortex*, bhu025.

Longmore, C. A., Liu, C. H., Young, A. W. (2008). Learning faces from photographs. Journal of Experimental Psychology: Human Perception and Performance, 34(1), 77−100.

Lundqvist, D., Flykt, A., Öhman, A. (1998). *The Karolinska Directed Emotional Faces – KDEF* [CD ROM]. Stolkholm: Karolinska Institutet, Department of Clinical Neuroscience, Psychology Section.

Markham, R., & Wang, L. (1996). Recognition of emotion by Chinese and Australian children. *Journal of Cross-Cultural Psychology, 27*(5), 616-643.

Matsumoto, D. (1989). Cultural influences on the perception of emotion. *Journal of Cross-Cultural Psychology, 20*, 92–105.

Matsumoto, D. (1992). American-Japanese cultural differences in the recognition of universal facial expressions. *Journal of Cross-cultural Psychology, 23*(1), 72−84.

Matsumoto, D. (2002). Methodological requirements to test a possible in-group advantage in judging emotions across cultures: comment on Elfenbein and Ambady (2002) and evidence. *Psychological Bulletin, 128*(2), 236−242.

Matsumoto, D., & Assar, M. (1992). The effects of language on judgments of universal facial expressions of emotion. *Journal of Nonverbal Behavior, 16*(2), 85−99.

Matsumoto, D., & Ekman, P. (1989). American-Japanese cultural differences in intensity ratings of facial expressions of emotion. *Motivation and Emotion, 13*(2), 143−157.

Mattavelli, G., Sormaz, M., Flack, T., Asghar, A. U., Fan, S., Frey, J., Manssuer, L., Usten, D., Young, A. W., & Andrews, T. J. (2013). Neural responses to facial expressions support the role of the amygdala in processing threat. *Social cognitive and affective neuroscience*, nst162.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6*(6), 255−260.

McAndrew, F. T. (1986). A cross-cultural study of recognition thresholds for facial expressions of emotion. *Journal of Cross-Cultural Psychology, 17*(2), 211−224.

McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of cognitive neuroscience, 9*(5), 605-610.

McConahay, J. B. (1986). Modern racism, ambivalence, and the modern racism scale. McConahay, J. B., Dovidio, J. F. (Ed); Gaertner, S. L. (Ed), (1986). *Prejudice, discrimination, and racism.* , (pp. 91-125). San Diego, CA: Academic Press.

Megreya, A. M., & Burton, M. (2006). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & Cognition, 34*(4), 865−876.

Megreya, A. M., White, D., & Burton, A. M. (2011). The other race effect does not rely on memory: evidence from a matching task. *The Quarterly Journal of Experimental Psychology, 64*, 1473–1483.

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

Michel, C., Caldara, R., & Rossion, B. (2006a). Same-race faces are perceived more holistically than other-race faces. *Visual Cognition, 14*(1), 55−73.

Michel, C., Corneille, O., & Rossion, B. (2007). Race categorization modulates holistic face encoding. *Cognitive Science, 31*, 911−924.

Michel, C., Rossion, B., Bülthoff, I., Hayward, W. G., & Vuong, Q. C. (2013). The contribution of shape and surface information in the other-race effect. *Visual Cognition, 21*, 9−10, 1202−1223.

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

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

Natu, V., Raboy, D., O'Toole, A. (2011). Neural correlates of own- and other-race face perception: spatial and temporal response differences. *NeuroImage, 54*, 2547−2555.

Niisbett, R. E., & Masuda, T. (2003). Culture and point of view. *Proceedings of the National Academy of Science, 100*(19), 11163−11170.

Nili, H., Wingfield, C., Walther, A., Su, L., Marslen-Wilson, W., & Kriegeskorte, N. (2014). A toolbox for representational similarity analysis. *PLoS Computational Biology, 10*(11), e1003915.

Oatley, K., & Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition & Emotion, 1*, 29−50.

Palermo, R., & Coltheart, M. (2004). Photographs of facial expression: Accuracy, response times, and ratings of intensity. *Behavior Research Methods, Instruments, & Computers, 36*(4), 634−636.

Pettigrew, T. F., & Tropp, L. (2000). Does intergroup contact reduce prejudice? Recent meta-analytic findings. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 93–114). New Jersey: Lawrence Erlbaum Associates.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage, 16*(2), 331−348.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of cognitive neuroscience, 12*(5), 729−738.

Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., & Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face-selective cortical regions. *NeuroImage, 56*(4), 2356−2363.

Pitcher, D., Duchaine, B., & Walsh, V. (2014). Combined TMS and fMRI reveal dissociable cortical pathways for dynamic and static face perception. *Current Biology, 24(17)*, 2066−2070.

Pitcher, D., Duchaine, B., Walsh, V., Yovel, G., & Kanwisher, N. (2011). The role of lateral occipital face and object areas in the face inversion effect. *Neuropsychologia, 49* (12), 3448−3453.

Puce, A., Allison, T., Asgari, M., Gore, J. C. & McCarthy, G. (1996) Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *Journal of Neuroscience, 16*, 5205–5215.

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *The Journal of Neuroscience, 18*(6), 2188-2199.

Rhodes, G., Hayward, W. G., & Winkler, C. (2006). Expert face coding: configural and component coding of own-race and other-race faces. *Psychonomic Bulletin & Review, 13*(3), 499−505.

Rhodes, G., Lie, H. C., Ewing, L., Evangelista, E., & Tanaka, J. W. (2010). Does perceived race affect discrimination and recognition of ambiguous-race faces? A test of the sociocognitive hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(1), 217−223.

Rice, G. E., Watson, D. M., Hartley, T., & Andrews, T. J. (2014). Low-level image properties of visual objects predict patterns of neural response across category-selective regions of the ventral visual pathway. *The Journal of Neuroscience, 34*(26), 8837−8844.

Richeson, J. A., Baird, A. A., Gordon, H. L., Heatherton, T. F., Wyland, C. L., Trawalter, S., & Shelton, J. N. (2003). An fMRI investigation of the impact of interracial contact on executive function. *Nature neuroscience, 6*(12), 1323−1328.

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011a). Beliefs alter holistic face processing… if response bias is not taken into account. *Journal of Vision, 11*(13), 17-17.

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011b). Holistic processing predicts face recognition. *Psychological Science, 22*(4), 464−471.

Ronquillo, J., Denson, T. F., Lickel, B., Lu, Z. L., Nandy, A., & Maddox, K. B. (2007). The effects of skin tone on race-related amygdala activity: An fMRI investigation. *Social cognitive and affective neuroscience, 2*(1), 39-44.

Rosenthal, R., Hall, J. A., DiMatteo, M. R., Rogers, P. L., & Archer, D. (1979). *Sensitivity to nonverbal communications: The PONS Test.* Baltimore, MD: The Johns Hopkins University Press.

Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition, 21*(2), 139−253.

Rossion, B., & Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *NeuroImage, 39*(4), 1959−1979.

Rossion, B., & Retter, T. L. (2015). Holistic face perception: Mind the gap! *Visual Cognition, 23*(3), 379−398.

Rotshtein, P., Henson, R. N., Treves, A., Driver, J., & Dolan, R. J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nature neuroscience, 8*(1), 107-113.

Rozin, P., Haidt, J., & McCauley, C. R. (1993). Disgust. In: M. Lewis M & J. M. Haviland (editors). *Handbook of emotions*. New York: Guilford Press, 575-594.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology. 39*(6), 1161–1178.

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

Sander, D., Grafman, J., Zalla, T. (2003). *Review of Neuroscience, 14*, 303–316.

Sandford, A., & Burton, A. M. (2014). Tolerance for distorted faces: Challenges to a configural processing account of familiar face recognition. *Cognition, 132*(3), 262–268.

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

Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: an fMRI study. *Cognitive Brain Research, 20*, 81−91.

Schweinberger, S. R., Huddy, V., & Burton, A. M. (2004). N250r: A face-selective brain reponse to stimulus repetition. *NeuroReport, 15*(9), 1501−1505.

Schweinberger, S. R., Pickering, E. C., Burton, A. M., & Kaufmann, J. M. (2002). Human brain potential correlates of repetition priming in face and name recognition. *Neuropsychologica, 40*(12), 2057−73.

Schweinberger, S. R., & Soukup, G. R. (1998). Asymmetric relationships among perceptions of facial identity, emotion, and facial speech. *Journal of Experimental Psychology: Human perception and performance, 24*(6), 1748−1765.

Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin, 100*(2), 139−156.

Sneddon, I., McKeown, G., McRorie, M., Vukicevic, T. (2011). Cross-cultural patterns in dynamic ratings of positive and negative natural emotional behavior. *PLoS ONE, 6*(2), e14679.

Soto, J. A. & Levenson, R. W. (2009). Emotion recognition across cultures: The influence of ethnicity on empathic accuracy and physiological linkage. *Emotion, 9*(6), 874−884.

Sormaz, M., Watson, D. M., Smith, W. A. P., Young, A. W., & Andrews, T. J. (2016). Modelling the perceptual similarity of facial expressions from image statistics and neural responses. *NeuroImage, 129*, 64−71.

Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law, 7*(1), 36−97.

Swallow, K. M., Braver, T. S., Snyder, A. Z., Speer, N. K., & Zacks, J. M. (2003). Reliability of functional localization using fMRI. *NeuroImage, 20*, 1561-1577.

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

Tanaka, J. W., Kaiser, M. D., Butler, S., & Le Grand, R. (2012). Mixed emotions: holistic and analytic perception of facial expressions. *Cognition And Emotion, 26*(6), 961−977.

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

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

Tanaka, J. W., Kaiser, M. D., Butler, S., & Le Grand, R. (2012). Mixed emotions: holistic and analytic perception of facial expressions. *Cognition And Emotion, 26*(6), 961−977.

Trautmann, S. A., Fehr, T., & Herrmann, M. (2009). Emotions in motion: Dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. *Brain Research, 1284*, 100−115.

Valentine, T. (1988). Upsiede-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology, 79*, 471−491.

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2011). Modulation of the fusiform face area following minimal exposure to motivationally relevant faces: evidence of in-group enhancement (not out-group disregard). *Journal of Cognitive Neuroscience, 23*(11), 3343-3354.

Vizioli, L., Rousselet, G. A., & Caldara, R. (2010). Neural repetition suppression to identity is abolished by other-race faces. *Proceedings of the National Academy of Sciences, 107*(46), 20081−20086.

Walker, P. M., & Hewstone, M. (2006). A perceptual discrimination investigation of the own-race effect and intergroup experience. *Applied Cognitive Psychology, 20*, 461–475.

Walker, P. M., Silvert, L., Hewstone, M., & Nobre, A. C. (2007). Social contact and other-race face processing in the human brain. *Social Cognitive and Affective Neuroscience, nsm035.*

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

Wang, Y., & Luo, Y. J. (2005). Standardization and assessment of college students' facial expression of emotion. *Chinese Journal of Clinical Psychology, 4*(13), 396−398.

Watson, Hartley, Andrews (2014). Patterns of neural response to visual scenes are linked to the low-level properties of the image. *NeuroImage, 99*, 402−410.

Watson, Hymers, Hartley, Andrews (2016). Patterns of neural response in scene-selective regions of the human brain are affected by low-level manipulations of spatial frequency. *NeuroImage, 124*, 107−117.

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.

Weibert, K. & Andrews, T. J. (2015). Activity in the right fusiform face area predicts the behavioural advantage for the perception of familiar faces. *Neuropsychologia, 75*, 588−596.

Wheeler, M. E., & Fiske, S. T. (2005). Controlling racial prejudice social-cognitive goals affect amygdala and stereotype activation. *Psychological Science, 16*(1), 56−63.

White, M. (2000). Parts and wholes in expression recognition. *Cognition and Emotion, 14*(1), 39−60.

Wiggers, M. (1982). Judgements of facial expressions of emotion predicted from facial behaviour. *Journal of Nonverbal Behavior, 7*(2), 101−116.

Williams, M. A., Morris, A. P., McGlone, F., Abbott, D. F., & Mattingley, J. B. (2004). Amygdala responses to fearful and happy facial expressions under conditions of binocular suppression. *Journal of Neuroscience, 24*, 2898–2904.

Winston, J. S., Henson, R. N. A., Fine-Goulden, M. R., & Dolan, R. J. (2004). fMRI-adaption reveals dissociated neural representations of identity and expression in face perception. *Journal of Neurophysiology, 92*, 1830–1839.

Winston, J. S., O'Doherty, J., & Dolan, R. J. (2003). Common and distinct neural responses during direct and incidental processing of multiple facial emotions. *NeuroImage, 20*(1), 84-97.

Wright, C. I., Martis, B., Schwartz, C. E., Shin, L. M., Fischer, H. H., McMullin, K., et al. (2003). Novelty responses and differential effects of order in the amygdala, substantia innominata, and inferior temporal cortex. *NeuroImage, 18*, 660–669.

Yan, X. Q., Andrews, T. J., & Young, A. W. (2016a). Cultural similarities and differences in perceiving and recognizing facial expressions of basic emotions. *Journal of Experimental Psychology: Human Perception and Performance, 42*(3), 423−440.

Yan, X. Q., Andrews, T. J., Young, A. W., & Jenkins, R. (2016b). Cross-cultural differences and similarities underlying other-race effects for facial identity and expression. *The Quarterly Journal of Experimental Psychology, 69*, 1247−1254.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*, 141−145.

Yoder, A. M., Widen, S. C., & Russell, J. A. (2016). The word disgust may refer to more than one emotion. *Emotion, 16*(3), 301−308.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configural information in face perception. *Perception, 16*, 747−759.

Young, S. G. & Hugenberg, K. (2010). Mere social categorization modulates identification of facial expressions of emotion. *Journal of Personality and Social Psychology, 99*(6), 964−977.

Young, A. W., McWeeny, K. H., Hay, D. C., & Ellis, A. W. (1986). Matching familiar and unfamiliar faces on identity and expression. *Psychological research, 48*(2), 63−68.

Young, A. W., Newcombe, F., de Haan, E. H., & Hay, D. C. (1993). Face perception after brain injury. *Brain, 116*(4), 941–959.

Young, A. W., Perrett, D. I., Calder, A. J., Sprengelmeyer, R., & Ekman, P. (2002). *Facial Expressions of Emotion: Stimuli and Tests (FEEST)*. Thames Valley Test Company, Bury St Edmunds, England.

Yovel, G., & Duchaine, B. C. (2006). Specialized face perception mechanisms extract both part and spacing information: Evidence from developmental prosopagnosia. *Journal of Cognitive Neuroscience, 18*, 580–593.

Yovel, G., & Kanwisher, N. (2004a). Face perception: domain specific, not process specific. *Neuron, 44*, 889–898.

Yovel, G., & Kanwisher, N. (2004b). Face perception engages a domain-specific system for processing both configural and part-based information about faces [Abastract]. *Journal of Vision, 4,* 133a.

Yovel, G., & Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Current Biology, 15*, 2256–2262.

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, F., Parmley, M., Wan, X. A., & Cavanagh, S. (2014). Cultural differences in recognition of subdued facial expressions of emotions. *Motivation and Emotion, 39*(2), 309−319.

Zhao, M. T., Hayward, W. G., & Bülthoff, I. (2014). Holistic processing, contact, and the other-race effect in face recognition. *Vision Research, 105*, 61–69.

Zhu, Z., Ho, S. M., & Bonanno, G. A. (2013). Cultural similarities and differences in the perception of emotional valance and intensity: A comparison of Americans and Hong Kong Chinese. *The American Journal of Psychology, 126*(3), 261−273.