Journal of Experimental Psychology:
Human Perception and Performance
2016, Vol. 42, No. 3, 423–440

0096-1523/16/$12.00 http://dx.doi.org/10.1037/xhp0000114

Cultural Similarities and Differences in Perceiving and Recognizing Facial Expressions of Basic Emotions

Xiaoqian Yan, Timothy J. Andrews, and Andrew W. Young
University of York

The ability to recognize facial expressions of basic emotions is often considered a universal human ability. However, recent studies have suggested that this commonality has been overestimated and that people from different cultures use different facial signals to represent expressions (Jack, Blais, Scheepers, Schyns, & Caldara, 2009; Jack, Caldara, & Schyns, 2012). We investigated this possibility by examining similarities and differences in the perception and categorization of facial expressions between Chinese and white British participants using whole-face and partial-face images. Our results showed no cultural difference in the patterns of perceptual similarity of expressions from whole-face images. When categorizing the same expressions, however, both British and Chinese participants were slightly more accurate with whole-face images of their own ethnic group. To further investigate potential strategy differences, we repeated the perceptual similarity and categorization tasks with presentation of only the upper or lower half of each face. Again, the perceptual similarity of facial expressions was similar between Chinese and British participants for both the upper and lower face regions. However, participants were slightly better at categorizing facial expressions of their own ethnic group for the lower face regions, indicating that the way in which culture shapes the categorization of facial expressions is largely driven by differences in information decoding from this part of the face.

Keywords: perception, categorization, facial expression, emotion, culture

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 20th century, theories of cultural relativism held sway, and it was thought that facial expressions were primarily culturally constructed and learned (see Elfenbein & Ambady, 2002b). However, for the last 40 years, the dominant position has been based on Ekman's (1980) interpretation of Darwin's (1872/1904) 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 (Ekman, 1972; Ekman, Sorenson, & Friesen, 1969; Izard, 1971). Studies have also revealed that this *own-group advantage* is lower for groups with a closer geographical distance or groups that have 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 and East Asian participants. In an eye-tracking study, Jack, Blais, Scheepers, Schyns, and Caldara (2009) found that when recognizing facial expressions, Western participants fixated on features in the eye and mouth regions, whereas East Asian participants mainly fixated on 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 nonetheless encoded, Jack, Caldara, and Schyns (2012) used a reverse correlation method to determine whether these differences in fix-

This article was published Online First October 19, 2015.

Xiaoqian Yan, Timothy J. Andrews, and Andrew W. Young, Department of Psychology, University of York.

This work was funded by an Overseas Research Scholarship (ORS) studentship to Xiaoqian Yan from the University of York. We are grateful to YueJia Luo for allowing us to use the set of Chinese facial expressions. We thank Claudia Chisari, Layal Husain, and Miaomiao Yu for helping us collect some of the data.

Correspondence concerning this article should be addressed to Xiaoqian Yan, Department of Psychology, University of York, UK. YO10 5DD. E-mail: xy760@york.ac.uk

ation 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 judgment 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 nonemotional faces (Blais, Jack, Scheepers, Fiset, & Caldara, 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 that 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: (a) a perceptual similarity task and (b) 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 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 that 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 Chinese participants and tested responses based on the whole face, the upper part of each face (eyes and forehead), or the lower part of each face (mouth and chin).

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 are 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 Facial Expressions of Emotion: Stimuli and Tests (FEEST) series (Young, Perrett, Calder, Sprengelmeyer, & Ekman, 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).

Method

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: (a) 20 Chinese students (mean age: 21.7 years old) and 20 British students (mean age: 19.5 years old), (b) nine Chinese students (mean age: 21.6 years old) and nine British students (mean age: 19 years old), and (c) 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 an additional four participants (two Chinese, two 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 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.

Stimuli. We used the same set of Ekman and Friesen (1976) faces selected from the FEEST set that have been used in recent studies (Harris, Young, & Andrews, 2012, 2014a, 2014b; Mattavelli et al., 2014). 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: (a) a high recognition rate for all expressions (mean recognition rate in a six-alternative forced-choice experiment: 93%), (b) consistency of the action units (muscle groups) across different individuals posing a particular expression, and (c) visual similarity of the posed expression across individuals. Using these criteria to select the individuals from the FEEST set helped 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° high and 7.5° wide.

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 completed 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 = *not very similar* and 7 = *very similar*. 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 15 different types of expression pairs (same-expression pairs: anger–anger, disgust–disgust, fear–fear, happiness–happiness, and 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 10 possible combinations for each of the five same-expression pairs, leading to a total of 50 same-expression pair trials. Similarly, there were 20 possible combinations for each of the 10 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:

1. 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 s, 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.

2. Perceptual similarity task with sequential presentation of two face stimuli. The two faces were presented successively in this task, separated by a 2-s fixation interval. The first face image was presented for 1.5 s and the second one for 2 s. This task was intended to be more difficult than Task 1 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.

3. Perceptual similarity task with sequential presentation and an interpolated mask (see example in Figure 1). In this task, the two face stimuli were presented sequentially, separated by a phase-scrambled face image mask for 1 s and also a 1-s fixation screen. The first face image was presented for 1.5 s and the second one for 2 s. 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 2 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.

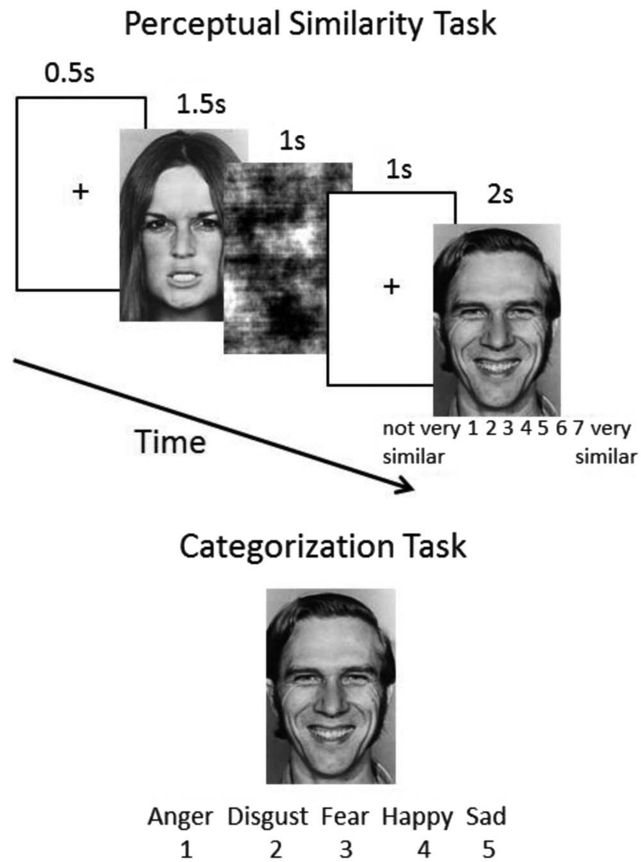

Figure 1. Examples of stimuli for the two tasks: perceptual similarity task (Variant 3) and categorization task in Experiment 1. All face images are from the Ekman and Friesen (1976) set selected from the FEEST series (Young, Perrett, Calder, Sprengelmeyer, & Ekman, 2002).

In the categorization task (see Figure 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, happiness, or sadness. 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 (five models posing five basic emotions) were presented twice each for each participant in a random order, yielding a total of 50 trials. Participants were also given a practice task of identifying 10 other facial expression pictures before the formal experiment.

Instructions for the Western participants were given in English. Instructions for the 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 meanings of these five emotion words immediately after they finished the categorization task (see Appendix A). 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 United Kingdom (UK; see Appendix B).

Results

To analyze 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 15 averaged ratings were then used to create a matrix reflecting the perceived similarity between expressions for participants from each cultural background (see Figure 2). This then allowed us to measure the overall concordance between the ratings of British and Chinese participants by corre-

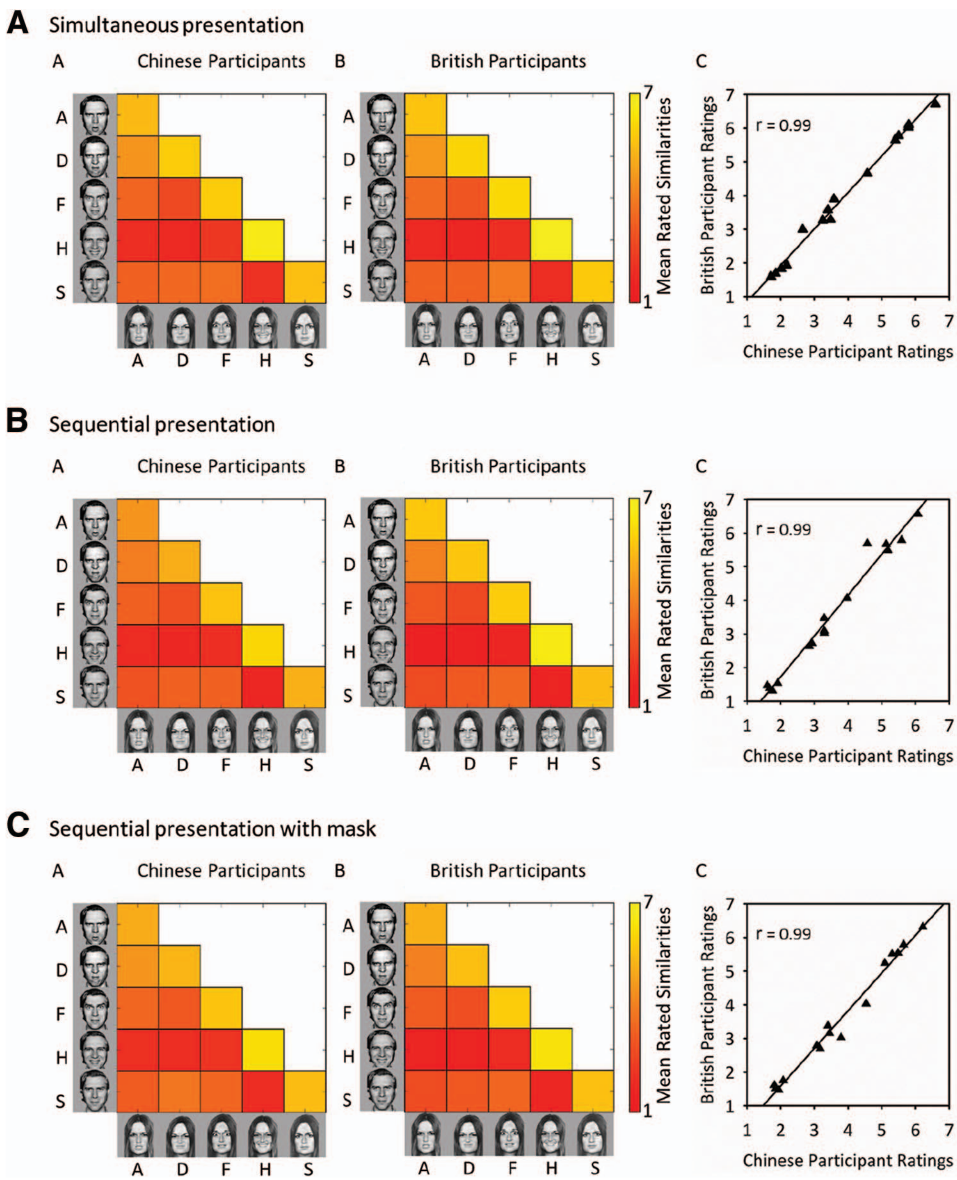

Figure 2. Correlation analyses of similarity ratings for three different versions of the perceptual similarity task (Variant 1—top row, simultaneous presentation; Variant 2—middle row, sequential presentation; Variant 3—bottom row, sequential presentation with an intervening mask) between Chinese and British participants in Experiment 1. Similarity matrices for Chinese (A) and British (B) participants (A = anger; D = disgust; F = fear; H = happiness; S = sadness). Scatterplot of rating correlations between the two groups of participants (C). All face images are from the Ekman and Friesen (1976) set selected from the FEEST series (Young et al., 2002). See the online article for the color version of this figure.

lating 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. These correlations were very high for each version of the task—$r = .99$, $p < .001$ for Variant 1, $r = .99$, $p < .001$ for Variant 2, and $r = .99$, $p < .001$ for Variant 3—indicating that the overall pattern of perception of the expressions between the two cultural groups was strikingly similar across all levels of task difficulty.

Although the aforementioned 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. We therefore recalculated the correlations with ratings of these same-expression pairs removed, leaving only ratings of the 10 combinations of different-expression pairs. In this way, we were able to estimate the structure of between-category differences themselves (e.g., 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: $r = .98$, $p < .001$ for Variant 1, $r = .99$, $p < .001$ for Variant 2, and $r = .98$, $p < .001$ for Variant 3. 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 performed the same categorization task, allowing for an overall analysis. A mixed analysis of variance (ANOVA) was conducted on the arcsine-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, $F(1, 76) = 8.13$, $p < .01$, $\eta^2 = 0.10$, 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 3. There were also main effects of expression, $F(4, 304) = 38.18$, $p < .001$, $\eta^2 = 0.33$, and a significant Expression × Group interaction, $F(4, 304) = 3.40$, $p = .01$, $\eta^2 = 0.04$. 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.

Although our principal focus of interest was on 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 analyzed 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, $F(1, 75) = 0.93$, $p > .1$, $\eta^2 = 0.12$. 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, $F(1, 75) = 0.93$, $p > .1$, $\eta^2 = 0.12$. These results indicate that

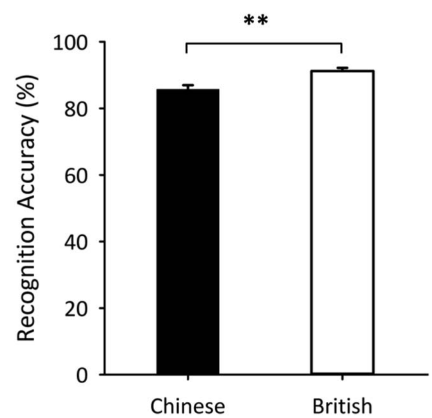

Figure 3. Overall emotion categorization accuracies (with standard error bars) for Chinese and British participants in Experiment 1. Asterisks denote higher overall emotion recognition rate for British participants in comparison with Chinese participants. ** $p < .01$.

there were no time differences for the participants in perceiving and identifying expressions of their own group 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 1 month to 9 years, but 27 out of 39 participants had been living in the UK for 1 year or less than 1 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 4A. The overall correlation was nonsignificant, $r = .07$, $p = .66$.

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, $r = .15$, $p = .36$ (see Figure 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.

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 (1976) 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

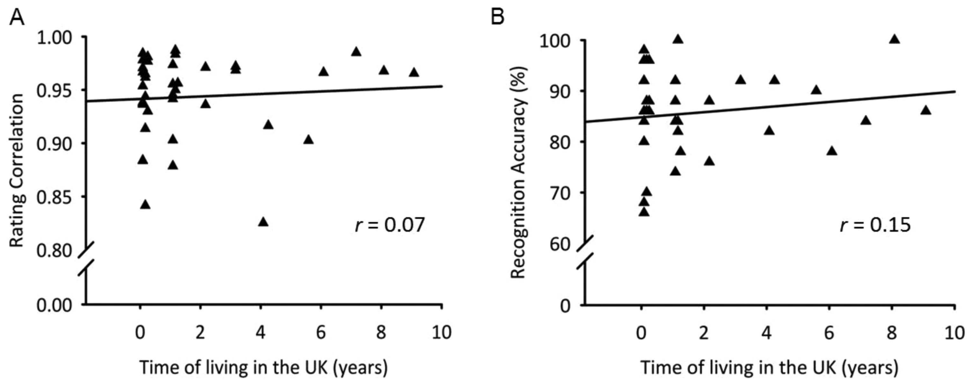

Figure 4. Scatterplots of Chinese participants' time living in the UK with their performance in the perceptual similarity task (A) and the categorization task (B) in Experiment 1.

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 (1976) series, this result is consistent with previous findings indicating the possible existence of an own-group advantage (Biehl et al., 1997; Ekman, 1972; Elfenbein & Ambady, 2002b; 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 colleagues (2012) suggested that East Asian and Western participants expect facial expressions to be primarily signaled 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.

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 (1976) faces to Chinese and British participants to further investigate potential strategy differences they might use in perceiving and categorizing facial expressions.

Method

Participants. Twenty Chinese (mean age: 21.5 years old) and 20 British (mean age: 20.3 years old) participants were recruited to perform 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 performed the same two tasks viewing only the lower half of faces. All participants were given a small a payment or course credit. In the lower region session, one Chinese participant's data were deleted mistakenly from the results, leaving only 19 participants in the Chinese group for analysis.

Stimuli. The images used in Experiment 1 were again used in this experiment, except that they were divided into upper and lower halves. The upper and lower halves of the faces were divided by a horizontal line through the middle of the bridge of the nose. The upper halves of faces were presented with a gray mask covering the lower part, and the lower halves of faces were presented with a gray mask covering the upper part (see Figure 5).

Procedure. Participants were asked to carry out Variant 3 of the perceptual similarity task (successive presentation with a mask) and the categorization task in this experiment. Apart from the use of only the upper half or lower half of images, all the other procedure details were identical to those in Experiment 1. As in Experiment 1, all Chinese participants showed correct understanding of the five emotion labels used in the categorization task.

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, $r = .99$, $p < .001$, and for the lower half, $r = .99$, $p < .001$. The average ratings for each group of participants and scatterplots are shown in Figure 6.

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 region and 0.99 ($p < .001$) for the lower region.

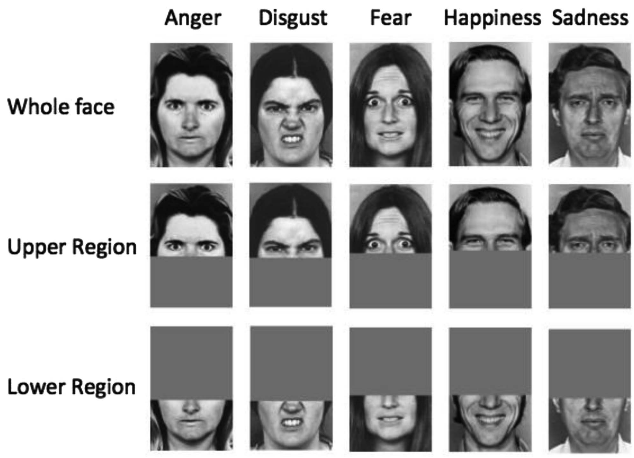

Figure 5. Examples of whole, upper half, and lower half face images for each expression posed by a different model from the FEEST set. Only the upper and lower region face images were used in Experiment 2. All face images are from the Ekman and Friesen (1976) set selected from the FEEST series (Young et al., 2002).

In the categorization task, as would be expected, recognition rates for emotion from partial 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 of the face: 0.75 (0.27), lower half of the 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 arcsine-transformed recognition accuracies found a significant Group x Face Region interaction, $F(1, 75) = 9.08$, $p < .01$, $\eta^2 = 0.11$ (see Figure 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 of faces, $t(38) = 1.33$, $p > .1$, whereas for the lower half of 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 (1976) faces, the partial expressions are posed by members of their own ethnic group for the British participants.

Main effects of expression, $F(4, 300) = 99.17$, $p < .001$, $\eta^2 = 0.57$, and face region, $F(1, 75) = 7.52$, $p < .01$, $\eta^2 = 0.09$, were also found, as well as a significant two-way Expression × Face Region interaction, $F(4, 300) = 54.6$, $p < .001$, $\eta^2 = 0.42$. 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 faces than the lower region, whereas disgust, $t(77) = 8.23$, $p < .001$, and happiness, $t(77) = 1.97$, $p = .05$, expressions 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 (1976) faces (Calder, Young, Keane, & Dean, 2000). The three-way Expression × Face Region × Group interaction was borderline but not significant, $F(4, 300) = 2.14$, $p = .08$, $\eta^2 = 0.03$. Other effects were not significant either.

Again, we analyzed 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 group 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 3 months to 6 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 regions of faces. For the upper region of faces, the correlations between the time Chinese participants had lived in the UK and their performance on 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 on 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.

Discussion

In Experiment 2, we investigated potential cultural differences in perceiving and categorizing facial expressions by showing only the upper or lower half of 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 (1976) set in much the same way as each other. Moreover, they appear to use information from the upper half 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 partial 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. Nonetheless, the results also indicated a possible own-group advantage for the lower half of faces. To be more specific, no reliable differences were found in recognition accuracy between the two cultural groups with the presentation of the upper half of 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) 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,

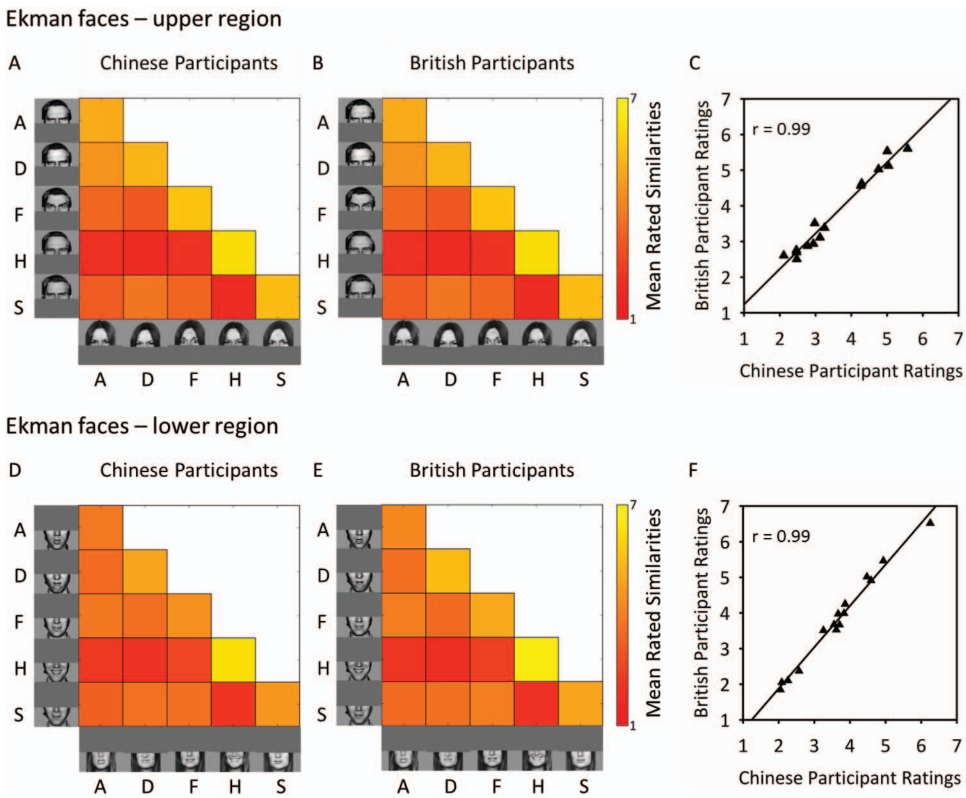

Figure 6. Correlation analyses of similarity ratings between Chinese and British participants in Experiment 2. Similarity matrices for Chinese (A) and British (B) participants for presentations of the upper face region (A = anger; D = disgust; F = fear; H = happiness; S = sadness). Scatterplot of rating correlations between the two groups for the upper region (C). Similarity matrices for Chinese (D) and British (E) participants for presentations of the lower face region. Scatterplot of rating correlations between the two groups for the lower region (F). All face images are from the Ekman and Friesen (1976) set selected from FEEST series (Young et al., 2002). See the online article for the color version of this figure.

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. First, 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). Second, the stimuli we used portrayed highly standardized facial expressions that were created according to Ekman and Friesen's (1978) Facial Action Coding System (FACS). 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.

Experiment 3

In Experiment 3, we used sets of faces portrayed by models from both Chinese and Western cultures. In contrast to the Ekman and Friesen (1976) stimuli, which are based on 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 to further explore the pattern of results from Experiment 1 by using varied expressions with both Chinese and Western White faces as stimuli.

Method

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 minimize 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.

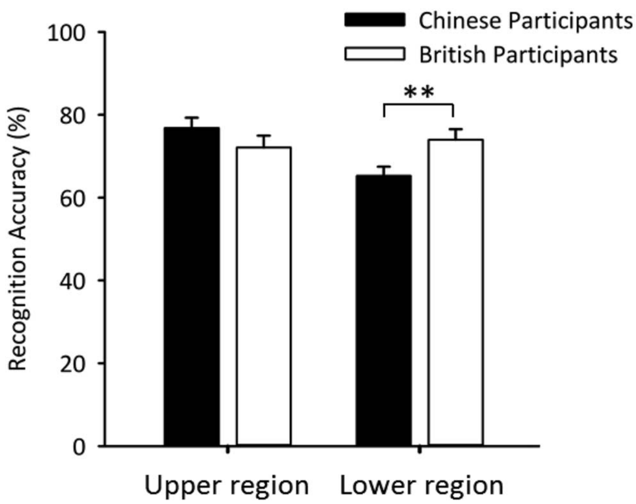

Figure 7. Overall emotion categorization accuracies (with standard error bars) for Chinese and British participants from upper and lower face regions in Experiment 2. Asterisks indicate the higher recognition rate for British participants in comparison with Chinese participants from the lower face region. ** $p < .01$.

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

There are 528 face images in total from five expression categories (anger, disgust, fear, happiness, and sadness) in the Chinese face set and 700 expression images with full-face poses 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 five 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 (six Chinese and six 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 five 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 (five 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 grayscale 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° high and 7° wide. Figure 8 shows examples of the images used in the following two experiments.

Procedure. Participants were required to perform either Variant 3 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 two blocks, yielding a total of four 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 s. The order of the emotion labels used for the categorization

Figure 8. Examples of whole, upper half, and lower half face images for each expression posed by a different model from the Chinese and Karolinska Directed Emotional Faces (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; Gong, Huang, Wang, & Luo, 2011; Wang & Luo, 2005), and all White faces are from the KDEF (Lundqvist, Flykt, & Öhman, 1998). The IDs of the KDEF images shown here are AM24ANS, AF09DIS, AF01AFS, AF08HAS, and AF26SAS, respectively.

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.

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, $r = .99$, $p < .001$, and also for the KDEF faces, $r = .98$, $p < .001$. The representational similarity matrices and scatterplots for the two groups of participants and the two sets of faces are shown in Figure 9.

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 arcsine-transformed recognition accuracies was conducted to analyze the categorization data. The most important result was a significant two-way Face × Group interaction, $F(1, 38) = 26.11$, $p < .001$, $\eta^2 = 0.41$. 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 10). This two-way interaction was also qualified by a

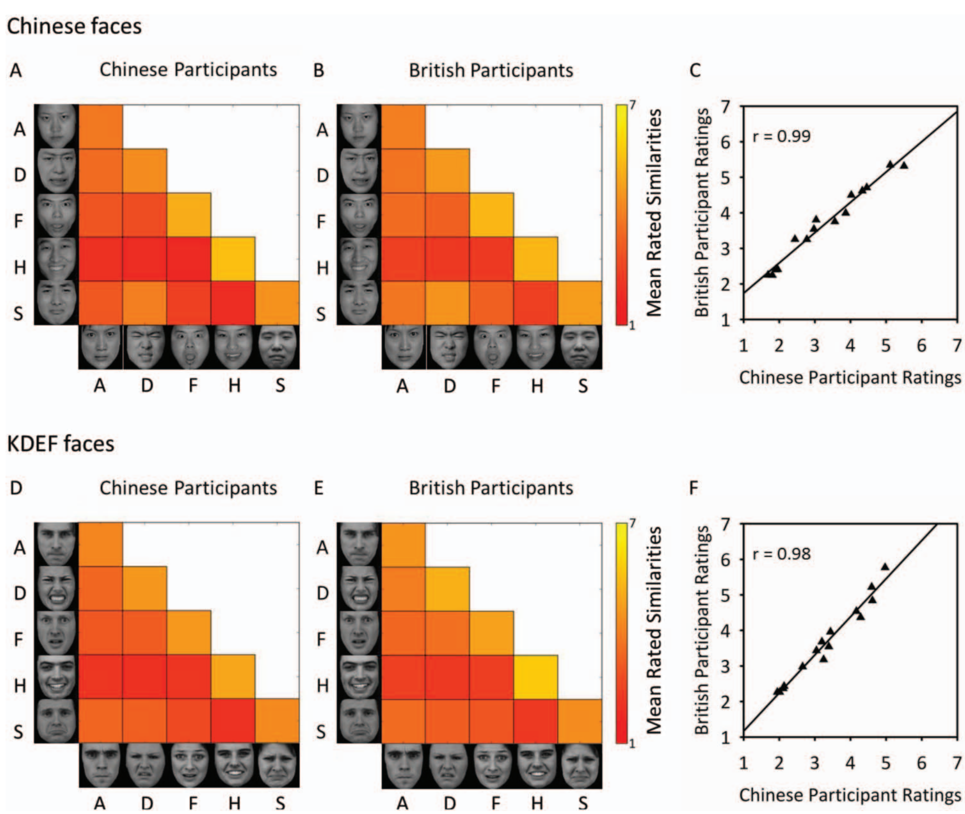

Figure 9. Correlation analyses of similarity ratings between Chinese and British participants in Experiment 3. Similarity matrices for Chinese (A) and British (B) participants for presentations of Chinese faces (A = anger; D = disgust; F = fear; H = happiness; S = sadness). Scatterplot of rating correlations between the two groups for Chinese faces (C). Similarity matrices for Chinese (D) and British (E) participants for presentations of Karolinska Directed Emotional Faces (KDEF). Scatterplot of rating correlations between the two groups for KDEF faces (F). All Chinese face images are from the Chinese Facial Affective Picture System (CFAPS; Gong et al., 2011; Wang & Luo, 2005), and all White faces are from the KDEF (Lundqvist et al., 1998). See the online article for the color version of this figure.

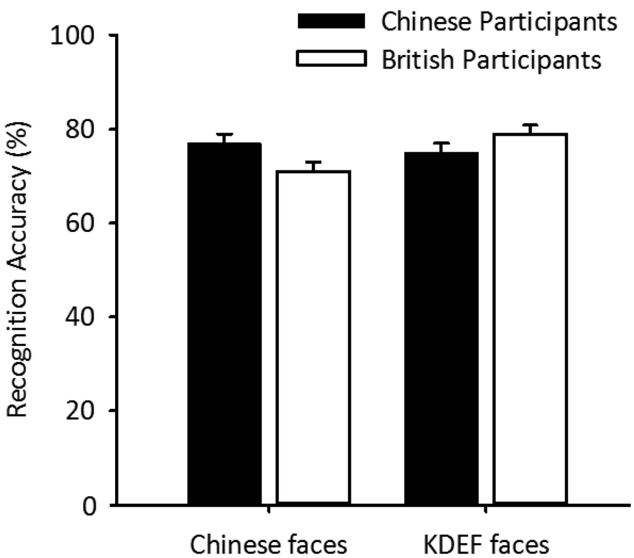

Figure 10. Overall recognition accuracies (with standard error bars) for Chinese and British participants from the Chinese faces and Karolinska Directed Emotional Faces (KDEF) in Experiment 3, plotting the statistically significant Group × Face interaction ($p < .001$).

three-way Expression × Face × Group interaction, $F(4, 152) = 7.02$, $p < .001$, $\eta^2 = 0.16$. 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, $F(4, 152) = 60.56$, $p < .001$, $\eta^2 = 0.61$, and face, $F(1, 38) = 8.25$, $p < .01$, $\eta^2 = 0.18$, as well as the Expression × Face interaction, $F(4, 152) = 39.65$, $p < .001$, $\eta^2 = 0.51$, 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.

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 group and the other group and also that 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 6 years and was less than 1 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.

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.

Experiment 4

Chinese and British participants carried out Variant 3 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.

Method

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

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 regions. The upper region faces were presented with a gray mask covering the lower part, and the lower region faces were presented with a gray mask covering the upper part.

Procedure. All participants had to complete the perceptual similarity task first and then completed the categorization task. In order to minimize the effect of task fatigue, the experiment was divided into two sessions. One half of participants performed 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 performed 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 (five models posing five 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 in the other experiments, all Chinese participants showed correct understanding of the meaning of the five emotion labels used in the categorization task.

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, $r = .98$, $p < .001$, and for the KDEF faces, $r = .96$, $p < .001$. 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, $r = .99$, $p < .001$, and for the KDEF faces, $r = .98$, $p < .001$. The average ratings for each group of participants and scatterplots are shown in Figure 11.

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 arcsine-transformed accuracy data found a significant Face × Region × Group interaction, $F(1, 76) = 8.53$, $p < .01$, $\eta^2 = 0.10$. 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 12).

We also found significant main effects of expression, $F(4, 304) = 118.53$, $p < .001$, $\eta^2 = 0.61$, face, $F(1, 76) = 64.15$, $p < .001$, $\eta^2 = 0.46$, and region, $F(1, 76) = 32.16$, $p < .001$, $\eta^2 = 0.30$. In addition to these main effects, the analysis also revealed significant two-way interactions of Expression × Region, $F(4, 304) = 8.81$, $p < .001$, $\eta^2 = 0.10$, Expression × Group, $F(4, 304) = 2.92$, $p < .05$, $\eta^2 = 0.04$, Expression × Face, $F(4, 304) = 34.15$, $p < .001$, $\eta^2 = 0.31$, Face × Group, $F(1, 76) = 19.37$, $p < .001$, $\eta^2 = 0.20$, and Face × Region, $F(1, 76) = 12.53$, $p < .001$, $\eta^2 = 0.14$. Three significant three-way interactions were also found: (a) Expression × Region × Group, $F(4, 304) = 2.55$, $p < .05$, $\eta^2 = 0.03$. Further analyses showed that for the upper region faces, Chinese participants were better at identifying happy expressions than British participants, $t(38) = -3.30$, $p < .01$, while for the lower region faces, British participants were better at identifying disgust expressions than Chinese participants, $t(38) = 2.71$, $p < .05$. (b) Expression × Face × Group, $F(4, 304) = 4.66$, $p < .001$, $\eta^2 = 0.06$. Further analyses showed that Chinese participants were better at recognizing Chinese anger expressions than British participants, $t(78) = 3.69$, $p < .001$. (c) Expression × Face × Region, $F(4, 304) = 3.27$, $p < .01$, $\eta^2 = 0.04$. Further analyses showed that for the upper region of 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 of 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, $F(4, 304) = 0.85$, $p > .1$, $\eta^2 = 0.01$. The rest of the effects were not significant either.

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

Figure 11 (opposite). Correlation analyses of similarity ratings between Chinese and British participants in Experiment 4. Similarity matrices for Chinese (A) and British (B) participants for presentations of Chinese upper face region (A = anger; D = disgust; F = fear; H = happiness; S = sadness). Scatterplot of rating correlations between the two groups for Chinese upper region images (C). Similarity matrices for Chinese (D) and British (E) participants for presentations of Karolinska Directed Emotional Faces (KDEF) upper face region images. Scatterplot of rating correlations between the two groups for KDEF upper face region images (F). Similarity matrices for Chinese (G) and British (H) participants for presentations of Chinese lower face region images. Scatterplot of rating correlations between the two groups for Chinese lower region images (I). Similarity matrices for Chinese (J) and British (K) participants for presentations of KDEF lower face region images. Scatterplot of rating correlations between the two groups for KDEF lower face region images (L). All Chinese face images are from the Chinese Facial Affective Picture System (CFAPS; Gong et al., 2011; Wang & Luo, 2005), and all White faces are from the KDEF (Lundqvist et al., 1998). See the online article for the color version of this figure.

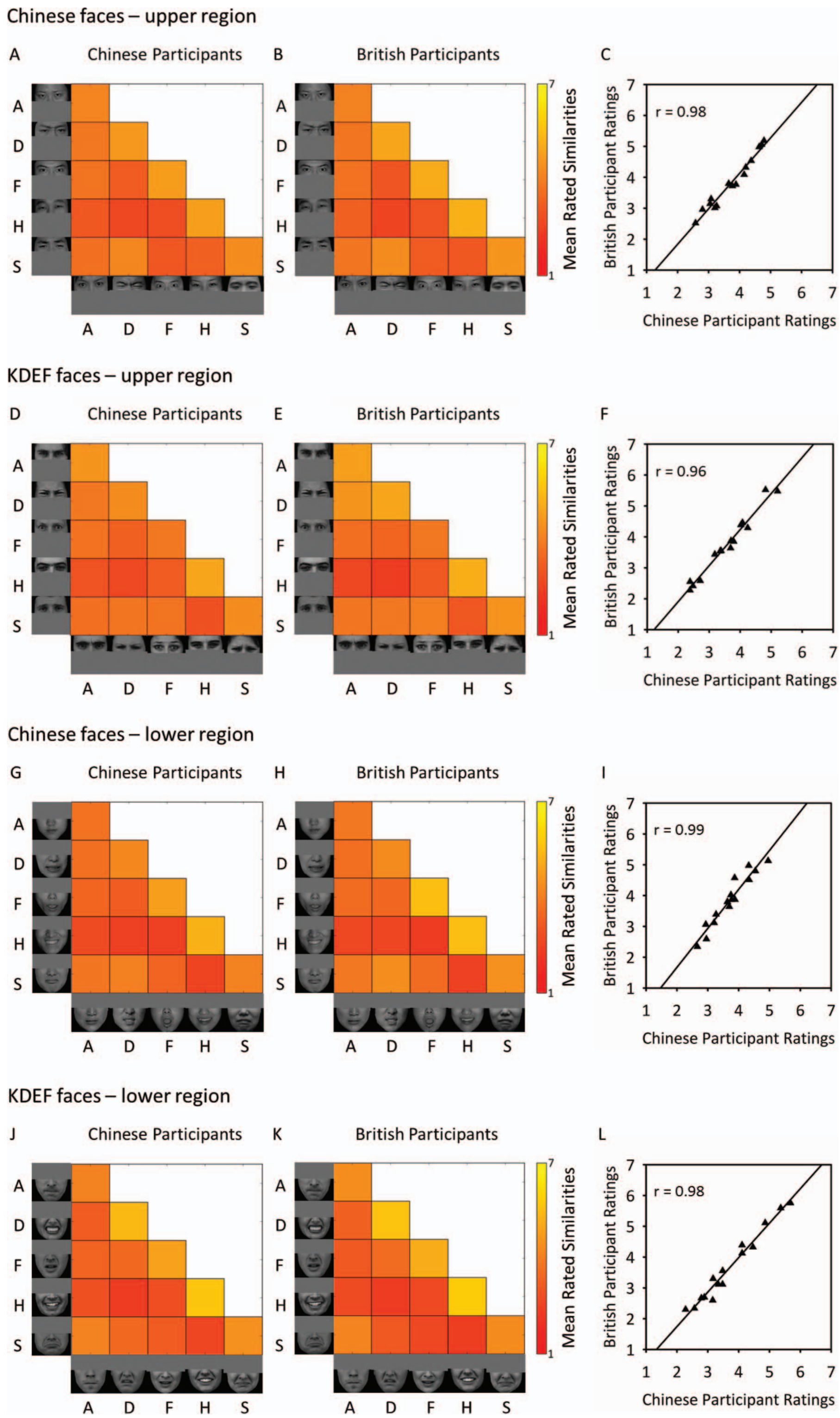

Figure 11 (opposite).

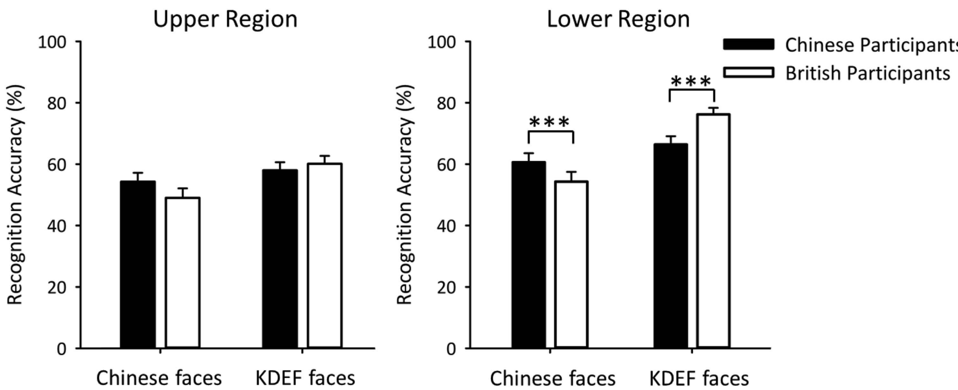

Figure 12. Overall emotion recognition accuracies (with standard error bars) for Chinese and British participants from upper and lower regions of the Chinese faces and Karolinska Directed Emotional Faces (KDEF) in Experiment 4. Asterisks denote conditions with a significantly higher recognition rate in comparison with the corresponding paired condition from the lower face region. *** $p < .001$.

were no time differences for participants in processing expressions of the own group and the other group and also that 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 6 years and was less than 1 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, $r = .01$, $p = .98$; 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, $r = .14$, $p = .55$; for KDEF faces, $r = .11$, $p = .65$; lower region categorization: for Chinese faces, $r = .12$, $p = .62$; for KDEF faces, $r = .02$, $p = .95$).

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 halves of 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 compared to Western faces and British participants can make better use of information from the mouth region of Western compared to Chinese faces.

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 differences in identity had to be ignored to make the perceptual judgment. In terms of perceiving facial expressions, we found no group differences in the patterns of interexpression similarity; correlations between Chinese and British participants for the rated perceptual similarities between pairs of expressions were always high across the 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 (Ekman, 1972; Ekman et al., 1969; Izard, 1971; 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, 2012). To address this question, we repeated the perceptual similarity and 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 the eyes or mouth).

The results from the categorization task with partial 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, which 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 that 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 group and the other group. Our results, however, indicated that participants from Chinese and British cultures spent the same amount of time on processing faces of either their own group or the other group. There was no evidence of speed–accuracy trade-offs or a general tendency to spend longer on evaluating own-group faces.

Neuroimaging studies have indicated that the posterior superior temporal sulcus (pSTS) and the amygdala respond to different types of changes 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 (1976) 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 conceptualizing 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 believe that 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 intercultural differences in categorization style (Nisbett & Masuda, 2003). Instead, consistent with the idea of universality (Darwin, 1872/1904; 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 United States for an average of 2.4 years were 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 the effect of time spent living 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): (a) Although the time our Chinese participants had been in the UK varied from 1 month to almost 9 years, many of them had been living in the UK for less than 1 year and (b) 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* (Hugenberg, Young, Bernstein, & Sacco, 2010; Meissner & Brigham, 2001; Pettigrew & Tropp, 2000). 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 and 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.

References

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

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

Bothwell, R. K., Brigham, J. C., & Malpass, R. S. (1989). Cross-racial identification. *Personality and Social Psychology Bulletin, 15,* 19–25. http://dx.doi.org/10.1177/0146167289151002

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

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,* 527–551. http://dx.doi.org/10.1037/0096-1523.26.2.527

Darwin, C. (1904). *The expression of the emotions in man and animals*. London, UK: Murray. (Original work published 1872) http://dx.doi.org/10.1037/10001-000

Du, S., Tao, Y., & Martinez, A. M. (2014). Compound facial expressions of emotion. *Proceedings of the National Academy of Sciences of the United States of America, 111,* E1454–E1462. http://dx.doi.org/10.1073/pnas.1322355111

Ekman, P. (1972). *Universals and cultural differences in facial expressions of emotions* (pp. 207–283). Lincoln: University of Nebraska Press.

Ekman, P. (1980). *The face of man: Expressions of universal emotions in a New Guinea village*. New York, NY: 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. http://dx.doi.org/10.1037/0033-2909.115.2.268

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behaviour: Categories, origins, usage, and coding. *Semiotica, 1,* 49–98. http://dx.doi.org/10.1515/semi.1969.1.1.49

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

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-cultural elements in facial displays of emotion. *Science, 164,* 86–88. http://dx.doi.org/10.1126/science.164.3875.86

Elfenbein, H. A., & Ambady, N. (2002a). Is there an in-group advantage in emotion recognition? *Psychological Bulletin, 128,* 243–249. http://dx.doi.org/10.1037/0033-2909.128.2.243

Elfenbein, H. A., & Ambady, N. (2002b). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin, 128,* 203–235. http://dx.doi.org/10.1037/0033-2909.128.2.203

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,* 92–110. http://dx.doi.org/10.1177/0022022102239157

Elfenbein, H. A., & Ambady, N. (2003b). Universals and cultural differences in recognizing emotions. *Current Directions in Psychological Science, 12,* 159–164. http://dx.doi.org/10.1111/1467-8721.01252

Elfenbein, H. A., & Ambady, N. (2003c). When familiarity breeds accuracy: Cultural exposure and facial emotion recognition. *Journal of Personality and Social Psychology, 85,* 276–290. http://dx.doi.org/10.1037/0022-3514.85.2.276

Freeman, J., & Ziemba, C. M. (2011). Unwrapping the ventral stream. *The Journal of Neuroscience, 31,* 2349–2351. http://dx.doi.org/10.1523/JNEUROSCI.6191-10.2011

Gong, X., Huang, Y. X., Wang, Y., & Luo, Y. J. (2011). Revision of the Chinese Facial Affective Picture System. *Chinese Mental Health Journal, 1,* 40–46.

Harris, R. J., 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 of the United States of America, 109,* 21164–21169. http://dx.doi.org/10.1073/pnas.1212207110

Harris, R. J., Young, A. W., & Andrews, T. J. (2014a). Brain regions involved in processing facial identity and expression are differentially selective for surface and edge information. *NeuroImage, 97,* 217–223. http://dx.doi.org/10.1016/j.neuroimage.2014.04.032

Harris, R. J., Young, A. W., & Andrews, T. J. (2014b). Dynamic stimuli demonstrate a categorical representation of facial expression in the amygdala. *Neuropsychologia, 56,* 47–52. http://dx.doi.org/10.1016/j.neuropsychologia.2014.01.005

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,* 1168–1187. http://dx.doi.org/10.1037/a0020463

Izard, C. E. (1971). *The face of emotion*. New York, NY: 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,* 1543–1548. http://dx.doi.org/10.1016/j.cub.2009.07.051

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,* 19–25. http://dx.doi.org/10.1037/a0023463

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

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,* 236–242. http://dx.doi.org/10.1037/0033-2909.128.2.236

Matsumoto, D., & Assar, M. (1992). The effects of language on judgments of universal facial expressions of emotion. *Journal of Nonverbal Behavior, 16,* 85–99. http://dx.doi.org/10.1007/BF00990324

Matsumoto, D., & Ekman, P. (1989). American-Japanese cultural differences in intensity ratings of facial expressions of emotion. *Motivation and Emotion, 13,* 143–157. http://dx.doi.org/10.1007/BF00992959

Mattavelli, G., Sormaz, M., Flack, T., Asghar, A. U., Fan, S., Frey, J., . . . Andrews, T. J. (2014). Neural responses to facial expressions support the role of the amygdala in processing threat. *Social Cognitive and Affective Neuroscience, 9,* 1684–1689. http://dx.doi.org/10.1093/scan/nst162

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

Nisbett, R. E., & Masuda, T. (2003). Culture and point of view. *Proceedings of the National Academy of Sciences of the United States of America, 100,* 11163–11170. http://dx.doi.org/10.1073/pnas.1934527100

Oatley, K., & Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition and Emotion, 1,* 29–50. http://dx.doi.org/10.1080/02699938708408362

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). Hilldale, NJ: Erlbaum.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39,* 1161–1178. http://dx.doi.org/10.1037/h0077714

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. http://dx.doi.org/10.1037/0033-2909.115.1.102

Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin, 100,* 139–156. http://dx.doi.org/10.1037/0033-2909.100.2.139

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

Young, A. W., Perrett, D. I., Calder, A. J., Sprengelmeyer, R., & Ekman, P. (2002). *Facial expressions of emotion: Stimuli and tests (FEEST).* Bury St Edmunds, UK: Thames Valley Test.

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. http://dx.doi.org/10.1016/j.jesp.2006.02.004

Appendix A

Questionnaire for English Labels Naming for Chinese Participants

中国被试英文情绪词命名问卷

请写出下面情绪单词的中文意思:
Anger ________________________
Disgust ________________________
Fear ________________________
Happy ________________________
Sad ________________________

Appendix B

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).

(*Appendices continue*)

....

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).

Received February 13, 2015
Revision received June 25, 2015
Accepted June 26, 2015 ■