Article:

Yan, Xiaoqian, Andrews, Tim orcid.org/0000-0001-8255-9120, Jenkins, Rob orcid.org/0000-0003-4793-0435 et al. (1 more author) (2016) Cross-cultural differences and similarities underlying other-race effects for facial identity and expression. Quarterly Journal of Experimental Psychology. pp. 1247-1254. ISSN: 1747-0226

https://doi.org/10.1080/17470218.2016.1146312

The Quarterly Journal of Experimental Psychology

Routledge
Taylor & Francis Group

ISSN: 1747-0218 (Print) 1747-0226 (Online) Journal homepage: http://www.tandfonline.com/loi/pqje20

Cross-Cultural Differences and Similarities Underlying Other-Race Effects for Facial Identity and Expression

Xiaoqian Yan, Timothy J. Andrews, Rob Jenkins & Andrew W. Young

Publisher: Taylor & Francis & The Experimental Psychology Society

Journal: *The Quarterly Journal of Experimental Psychology*

DOI: 10.1080/17470218.2016.1146312

CROSS-CULTURAL DIFFERENCES AND SIMILARITIES UNDERLYING OTHER-RACE EFFECTS FOR FACIAL IDENTITY AND EXPRESSION

Xiaoqian Yan, Timothy J. Andrews, Rob Jenkins and Andrew W. Young

Department of Psychology, University of York, York YO10 5DD United Kingdom

*Corresponding author: Xiaoqian Yan (xy760@york.ac.uk)

Figures: 4

Tables: 1

Word Count: 2998

Key words: face, race, identity, expression

Running head: ORE in facial identity and expression

Acknowledgements: XY was supported by an ORS studentship from the University of York.

We are grateful to YueJia Luo for allowing us to use the set of Chinese facial expressions.

ABSTRACT

Perceptual advantages for own-race compared to other-race faces have been demonstrated for the recognition of facial identity and expression. However, these effects have not been investigated in the same study with measures that can determine the extent of cross-cultural agreement as well as differences. To address this issue, we used a photo sorting task in which Chinese and Caucasian participants were asked to sort photographs of Chinese or Caucasian faces by identity or by expression. This paradigm matched the task demands of identity and expression recognition and avoided constrained forced-choice or verbal labelling requirements. Other-race effects of comparable magnitude were found across the identity and expression tasks. Caucasian participants made more confusion errors for the identities and expressions of Chinese than Caucasian faces, while Chinese participants made more confusion errors for the identities and expressions of Caucasian than Chinese faces. However, analyses of the patterns of responses across groups of participants revealed a considerable amount of underlying cross-cultural agreement. These findings suggest that widely repeated claims that members of other cultures "all look the same" overstate the cultural differences.

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, in press).

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 because 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 Caucasian backgrounds sorted both Chinese and Caucasian faces. We predicted that Chinese participants would make more confusion errors for Caucasian faces, while Caucasian participants would make more confusion errors 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., in press). 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 Caucasian participants, offering a complementary perspective on Yan et al.'s (in press) findings.

METHOD

Participants

Twenty Chinese students brought up in mainland China with Chinese parents (mean age, 22.6 years) and 20 Caucasian students brought up in western countries with 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.

Stimuli

Two sets of 20 Caucasian and two sets of 20 Chinese faces were created for the identity sorting task, and two sets of 20 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 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. (in press); 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 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.

Procedure

Each participant was asked to complete the sorting task for the 8 different sets of 20 stimuli; 2 Chinese Identity sets, 2 Caucasian Identity sets, 2 Chinese Expression sets, and 2 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.

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), confusion errors (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 confusion errors in which different people are mistaken for the same person. Following Jenkins et al. (2011), confusion errors 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 error score for each stimulus set. Performance for the two sets used for each sorting task (Chinese Identity, Caucasian Identity, Chinese Expression and Caucasian Expression) was then averaged for each participant.

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

Figure 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 ($F(1,38) = 37.86$, $MSE = 0.93$, $p < .001$, partial $\eta^2 = 0.50$).

Further simple effects analysis showed that there were significant differences in the confusion errors made by Chinese participants between Chinese and Caucasian faces ($F(1,38) = 12.94$, $MSE = 0.93$, $p < .001$), and in the confusion errors made by Caucasian participants between Chinese and Caucasian faces ($F(1,38) = 26.06$, $MSE = 0.93$, $p < .001$), 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, $F(1,38) = 0.71$, $MSE = 1.06$, $p = .4$, 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, $F(1,38) = 5.22$, $MSE = 2.76$, $p < .05$, partial $\eta^2 = 0.12$, reflecting a borderline difference between the number of confusions made by Chinese than by Caucasian participants in the expression task, $F(1,38) = 3.96$, $MSE = 2.13$, $p = .05$. No other significant effects were found.

Figure 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 ($F(1,38) = 58.13$, $MSE = 6.99$, $p < .001$, partial $\eta^2 = 0.61$). There was no effect of Face Ethnicity ($F(1,38) = 3.20$, $MSE = 0.78$, $p = .08$, partial $\eta^2 = 0.08$) or Participant Group ($F(1,38) = 0.78$, $MSE = 12.91$, $p > .1$, partial $\eta^2 = 0.02$). However, there was a significant interaction between Face Ethnicity and Participant Group ($F(1,38) = 9.26$, $MSE = 0.78$, $p < .01$, partial $\eta^2 = 0.20$). This was because there was a significant difference in the number of piles made by Caucasian participants for Chinese faces compared to Caucasian faces ($F(1,38) = 11.68$, $MSE = 0.78$, $p < .01$), whereas no reliable difference was observed between the number of piles made for Caucasian and Asian faces by Chinese participants ($F(1,38) = 0.78$, $MSE = 0.78$, $p > .01$). There were no other significant effects.

Figure 3 shows the sorting time for the identity and expression tasks. The ANOVA found a significant main effect of Face ($F(1,38) = 4.7$, $MSE = 0.93$, $p < .05$, partial $\eta^2 = 0.11$), with slightly more time spent on the Chinese faces compared with the Caucasian faces (Chinese faces: 3.9 min; Caucasian faces: 3.5 min). There was also a significant main effect of Task ($F(1,38) = 63.73$, $MSE = 2.34$, $p < .001$, partial $\eta^2 = 0.63$), 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.

The main finding from these analyses, then, was the Face Ethnicity x Participant Group interaction for confusion errors shown in Figure 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 shows examples of the response matrices for the groups of participants in one Caucasian identity and one Chinese expression sorting task.

From these response matrices we calculated a measure of cross-cultural agreement based on the overall correlations between the response matrices of Chinese and 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 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.

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, 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.

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. Caucasian participants made more confusion errors for the identities and expressions of Chinese than Caucasian faces, while Chinese participants made more confusions for the identities and expressions of 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 confusion errors. 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 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 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., in press), 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.

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.

REFERENCES

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

Chance, J. E., & Goldstein, A. G. (1996). The other race effect and eyewitness identification. In S. L. Sporer, R. S. Malpass, & G. Koehnken (Eds.), *Psychological Issues in Eyewitness Identification* (pp. 153−176). Mahwah, NJ: Erlbaum.

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

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.

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.

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

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.

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

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.

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.

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

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*(8), 1473−1483.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin, 115*(1), 102–141.

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.

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.

Yan, X. Q., Andrews, T. J., & Young, A. W. Cultural similarities and differences in perceiving and recognizing facial expressions of basic emotions. *Journal of Experimental Psychology: Human Perception and Performance* (In Press).

Figure Captions

Figure 1 Mean confusion errors (derived from piles containing more than one identity or more than one emotional expression) for Chinese and Caucasian participants in facial identity and expression sorting tasks involving Chinese and Caucasian faces (with standard error bars).

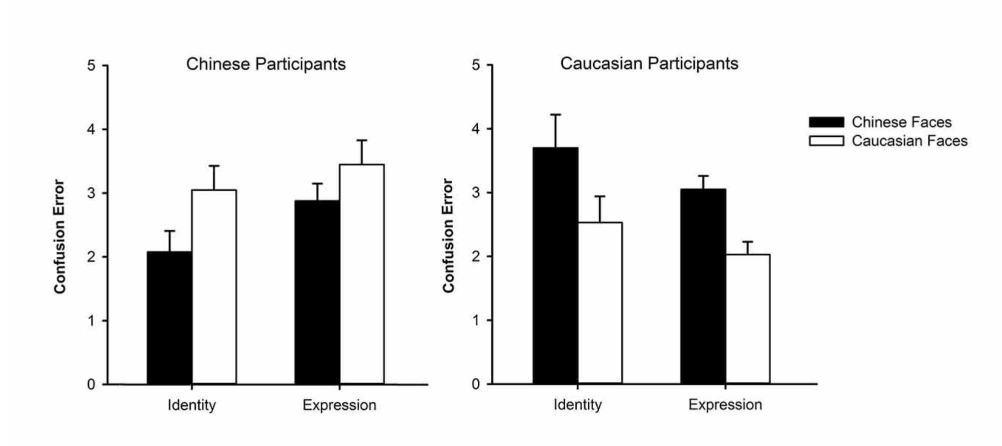
Figure 2 Mean numbers of piles created by Chinese and Caucasian participants in facial identity and expression sorting tasks involving Chinese and Caucasian faces (with standard error bars).

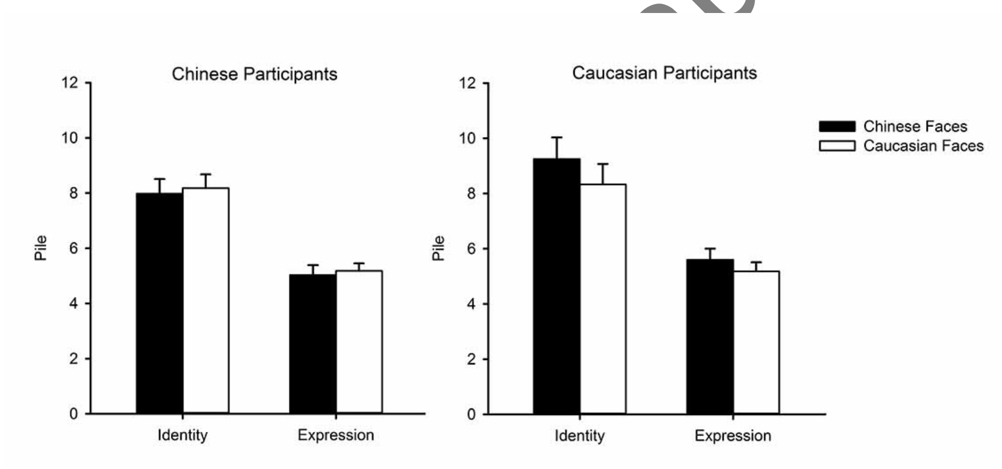
Figure 3 Mean sorting time (in minutes) for sets of 20 stimuli by Chinese and Caucasian participants in facial identity and expression sorting tasks involving Chinese and Caucasian faces (with standard error bars).

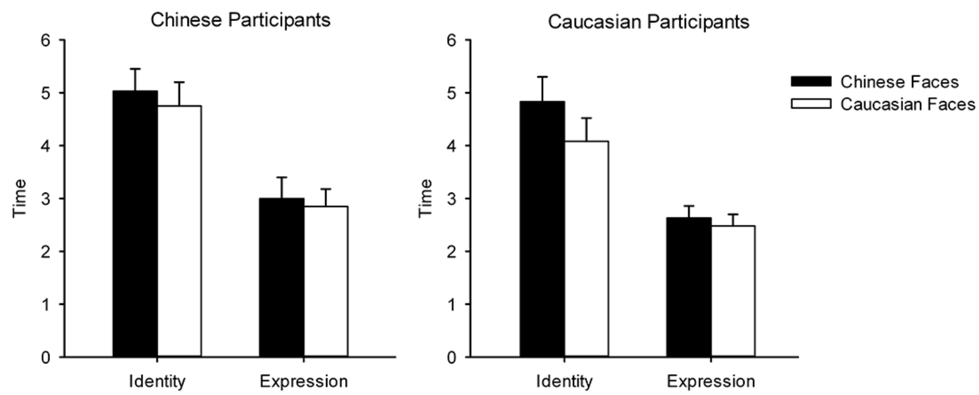
Figure 4 Response matrices for Chinese and Caucasian participants for one Caucasian Identity set (A) and one Chinese Expression set (B). The X- and Y-axes indicate the 5 different images of each of 4 identities/expressions. Each cell in the matrix represents the number of times that two images were sorted 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 Caucasian participants in both cases were 0.90, $p < .001$.

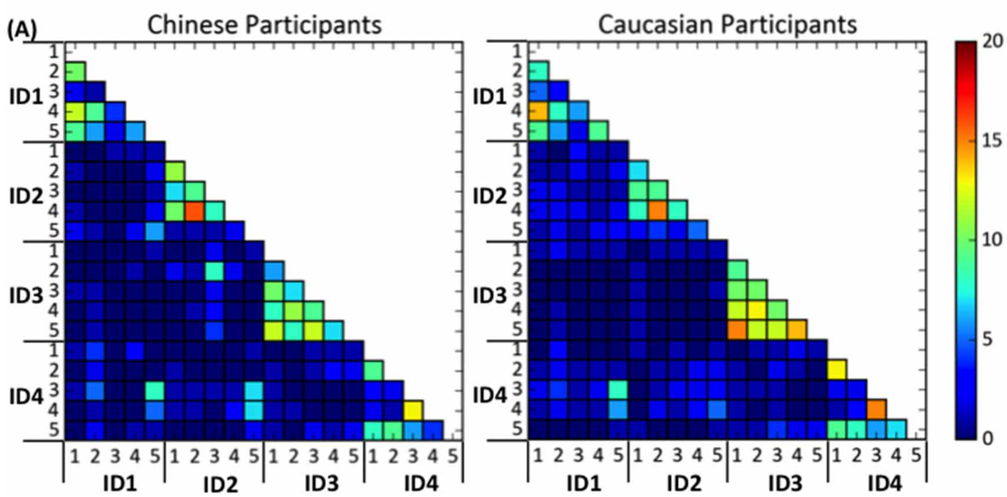
Table 1 Correlations between the sorting solutions of Chinese and Caucasian participants for the eight different sorting tasks. The Overall correlations use all the data from the corresponding response matrices (as shown in Figure 4). The Within correlations use only those cells in each matrix where 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.

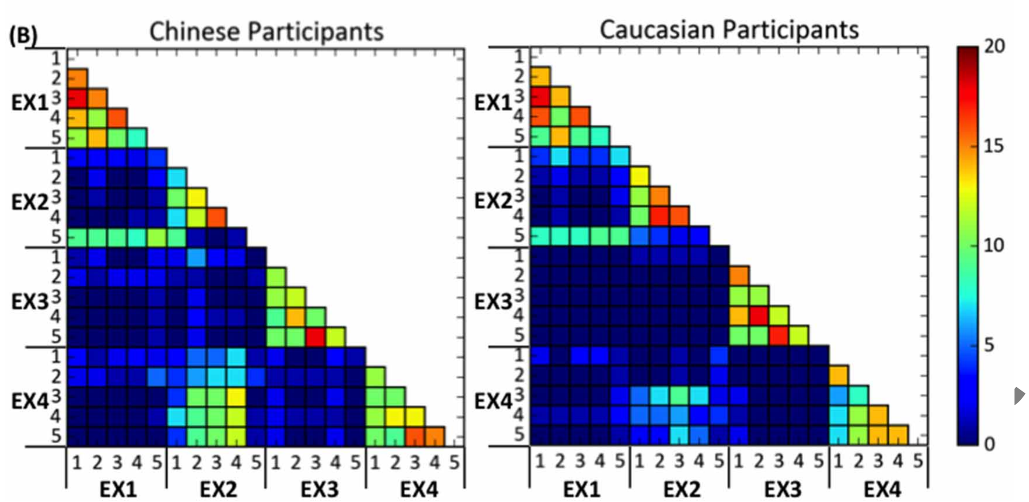
		Chinese Identity	Caucasian Identity	Chinese Expression	Caucasian Expression
Overall	Set1	0.70	0.85	0.90	0.90
	Set 2	0.84	0.90	0.80	0.91
Within	Set1	0.81	0.83	0.81	0.80
	Set 2	0.74	0.86	0.57	0.70
Between	Set1	0.51	0.47	0.75	0.73
	Set 2	0.36	0.57	0.61	0.62


Chinese Participants
Caucasian Participants
Confusion Error
Identity
Expression
Chinese Faces
Caucasian Faces


Chinese Participants
Caucasian Participants
Pile
Identity
Expression
Chinese Faces
Caucasian Faces


Chinese Participants
Caucasian Participants
Time
Identity
Expression
Chinese Faces
Caucasian Faces


(A)
Chinese Participants
Caucasian Participants
ID1
ID2
ID3
ID4


(B)
Chinese Participants
Caucasian Participants
EX1
EX2
EX3
EX4
1 2 3 4 5
0
5
10
15
20