Culture Variation in the Average Identity Extraction: The Role of Global vs. Local Processing Orientation

Shenli Peng[1], Chang Hong Liu[2], Xiaofan Yang[1], Haojian Li, Wenfeng Chen[1] & Ping Hu[1]

[1]Department of Psychology, Renmin University of China, Beijing, China
[2]Department of Psychology, Bournemouth University, Poole, UK
[3]School of Management and Economics, Chinese University of Hongkong, Shenzhen, Shenzhen, China

Corresponding concerning this article should be addressed to Ping Hu, Department of Psychology, Renmin University of China;
Address: No.59 Zhongguancun Avenue, Haidian District, Beijing, China;
E-mail: huping@ruc.edu.cn

Abstract

Research has shown that observers often spontaneously extract a mean representation from multiple faces/objects in a scene even when this is not required by the task. This phenomenon, now known as ensemble coding, has so far mainly been based on data from Western populations. This study compared East Asian and Western participants in an implicit ensemble-coding task, where the explicit task was to judge whether a test face was present in a briefly exposed set of faces. Although both groups showed a tendency to mistake an average of the presented faces as target, thus confirming the universality of ensemble coding, East Asian participants displayed a higher averaging tendency relative to the Westerners. To further examine how a cultural default can be adapted to global or local processing demand, our second experiment tested the effects of priming global or local processing orientation on ensemble coding via a Navon task procedure. Results revealed a reduced tendency for ensemble coding following the priming of local processing orientation. Together, these results suggest that culture can influence the proneness to ensemble coding, and the default cultural mode is malleable to a temporary processing demand.

Keywords: ensemble coding, cultural difference, global processing orientation, local processing orientation

Introduction

There is now substantial evidence that the visual system employs ensemble coding to rapidly summarize the gist information from a set of stimuli (Alvarez, 2011; Whitney & Leib, 2018). This has been demonstrated at both low-level and higher-level vision (Ariely, 2001; Haberman & Whitney, 2007, 2009). At high-level vision, participants could extract a mean representation from a set of faces after briefly viewing them on the screen. Participants tended to mistake a morphed average of the set as a set member even though it was never presented (de Fockert & Wolfenstein, 2009). The finding has since been replicated by numerous studies (Kramer et al., 2015; Neumann et al., 2013, 2018; Rhodes et al., 2015, 2018), which all reported comparable proportion of "present" responses to both set average and set member of the briefly viewed faces. The task has been referred to as "(implicit) ensemble coding paradigm" (Neumann et al., 2018), because the explicit task was to identify a set member rather than the set average. The fact that the average of an observed set is more likely to be chosen as a set member than the average of an unobserved set suggests that observers automatically extracted the average from the set of observed faces.

However, current knowledge on ensemble coding has mainly been derived from performance of Western participants. Although a small number of studies have also used Asian participants (Ji et al., 2014; Li et al., 2016; Chen & Zhou, 2018), few have compared their results across cultures. As suggested by Im et al. (2017), the sociocultural aspects of the ensemble coding of faces could be important, because there are reasons to assess the extent to which current findings on ensemble coding are generalizable across cultures. Few others in the existing literature have raised this question. The current literature produces the impression that the findings are universal. Indeed, research appears to show that participants from both East Asian and Western cultures are able to rapidly extract the ensemble information from a group of stimuli (Haberman & Whitney, 2007; Ji et al, 2014). This means there is a qualitative resemblance across cultural groups. However, are people in different cultures equally predisposed to ensemble coding? Can the degree of propensity for ensemble coding vary cross-culturally? Such issues cannot be answered without a direct comparison of ensemble coding performances across different cultural groups. A main purpose of this study, therefore, was to address the possibility of cultural influence on ensemble coding.

Why might culture affect ensemble coding? The conjecture is based on mounting evidence that the way we process visual information can be shaped by culture (Nisbett et al., 2001;

Miyamoto, 2013; Nisbett & Miyamoto, 2005, see Han et al., 2013 for a review). East Asians such as Chinese and Japanese tend to focus more on the holistic and global aspects of the visual world, whereas Westerners such as Americans and British tend to focus more on local information (e.g. Masuda & Nisbett, 2001). For example, the Masuda and Nisbett's study shows that after viewing an underwater scene of fish and other objects, Americans tended to remember the more salient fish on the foreground, whereas Japanese also recall items on the backgrounds, such as seaweeds. This suggests that Americans tend to pay more attention on focal information, whereas Japanese pay more attention to the relational information within the scene. This global vs. local processing bias is often attributed to some underlying cultural differences between individualism and collectivism.

Recent evidence for discrepant focus on global vs. local information comes from a study by Lao et al. (2013), who recorded the electrophysiological signals of Chinese and British observers during a matching task. Observers in the task judged whether a target was identical to the adaptor Navon figure. The hierarchical Navon figures (Navon, 1977) in their study contained five shapes, comprising of circle, cross, diamond, square and triangle. These shapes could be either shown at a global or local level. For example, a global square shape could consist of local squares or local circles. The target was congruent with the adaptor either globally or locally. Indexed by an early P1 component, Chinese observers displayed greater sensitivity to global congruency relative to British observers. This was consistent with the behavioral advantage of Chinese observers in the global task. Together with numerous other studies (e.g., Kitayama & Murata, 2013; Nisbett & Miyamoto, 2005), their results demonstrate a cultural difference in processing orientation, where Easterners tend to adopt a more global orientation, whereas Westerners display a more local orientation.

Such a cultural difference could play a role in ensemble coding, because an ensemble requires a global orientation. The idea that global processing is conducive to ensemble perception was firstly proposed by Chong and Treisman (2003), and further studied in Chong and Treisman, (2005a, 2005b), Chong et al. (2008) and Peng et al. (2019). Additionally, patients with unilateral spatial neglect (Pavlovskaya et al., 2015) and congenital prosopagnosia (Robson et al., 2018) displayed deficiency in ensemble perception, which might due to their focused attention mechanism and related perception of local features (local processing orientation). For example, Robson et al (2018) suggested that the weaker ensemble perception in participants with congenital

prosopagnosia relative to controls could be attributed to the impaired global processing, since Avidan et al. (2011) has demonstrated the impaired global processing in congenital prosopagnosia. Although the existing evidence leaves little question that both Western and East Asian observers extract a set average by focusing on the global properties of a scene, it remains possible observers from different cultures display varying levels of focus on the global information, which could in turn affect the level of propensity for ensemble coding. To date, few studies have tested this hypothesis. An exception to this is the recent study by Im et al. (2017), who presented Korean and American participants with two sets of emotional faces on the two sides of a screen. Both groups were instructed to perform an avoidance task based on the mean emotion of the two face sets. Even though the mean emotion of either set was not given, participants from both cultural groups displayed above-chance accuracy, suggesting that both Korean and American participants were able to extract a mean emotion from each set of faces. A more novel finding in their study was that the Korean participants showed higher overall accuracy than the American participants. The authors attributed this effect to different levels of sensitivity to global information between East Asians and Westerners. Furthermore, due to the relative strengths of collectivism and individualism, East Asians may see emotions in the sets as a group feeling while Westerners may see them as a collection of individuals' feelings (Masuda et al., 2008).

Despite this tantalizing explanation, a potential limitation of Im et al.'s avoidant task is that Asians may be more susceptible to attentional bias for avoidance-oriented information, hence their results could be attributed to the difference in the attentional bias rather than in sensitivity to set average. Some studies have reported different levels of sensitivity to avoidance-orientation information between Easterners and Westerners. In Elliot et al. (2001), for example, Asian Americans reported more avoidance goals than Caucasian Americans. This was consistent with the finding by Hamamura et al. (2009), who reported that North Americans were better at recall memory for approach-focused information than for avoidance-focused information, whereas Japanese displayed comparable performance for the two types of information. These studies thus suggest a cultural difference in Easterners and Westerners' sensitivity to approach-avoidance information. Hence, the finding of Im et al could also be attributed to the better ability of Koreans in processing avoidant task than North Americans. To rule out this alternative explanation of their effects based on ensemble performance, it is necessary to ascertain whether East Asians would

display a stronger ensemble coding bias without being influenced by avoidance-focused information. To this end, we adopted the identity recognition task from de Fockert and Wolfenstein (2009) in our Experiment 1, which compared the tendency of East Asian and Western participants to mistake a set average as a set member for faces that showed a neutral facial expression.

Apart from the potential cultural influence on ensemble coding, we were also interested in how malleable the cultural norm was for shifting attention between global and local information. Cultural studies have often shown both stable and malleable aspects of cultural norms. For example, in general Westerners tend to endorse an independent self-construal, seeing themselves as being independent, autonomous, and separate from others, whereas East Asians favor a more interdependent self-construal, emphasizing the importance of interconnectedness with others (Triandis, 1989; Markus & Kitayama, 1991; Ji et al., 2000; Lewis et al., 2008; Han et al., 2013). However, these cultural norms are also fairly flexible when individuals are temporarily exposed to other cultural frameworks. For example, although recognition of one's own face and friends' faces by Chinese participants are affected by their default cultural mode of interdependent self-construal (Sui et al., 2009), their responses can be shifted toward greater emphasis on self-face after exposure to the independent self-construal in a priming procedure (Sui & Han, 2007; Sui et al., 2013). Such evidence suggests that face perception is modulated both by culture, and by momentary adjustment within the cultural framework to fit a specific context. It is possible that ensemble coding is also adjustable according to a specific priming context. Assessing how ensemble coding may be affected by priming attention to global or local information was the goal of our Experiment 2.

Therefore, the present study aimed to explore cultural differences in the default mode of ensemble coding as well as the extent that this default mode of cultural bias could be adjusted or temporarily adapted to an encouraged processing orientation. To achieve these goals, we conducted two experiments, which both employed the task in de Fockert and Wolfenstein (2009). Participants were presented with a set of four faces and were subsequently asked to judge whether or not a test face had been present in the set. The first experiment compared the likelihood of East Asian and Western participants to judge the test face as being present in the set if it was the morphed mean of all four set faces than when it was an actual set member. The experiment tested

the hypothesis that due to a more holistic processing orientation, East Asian participants should display a stronger tendency to extract the set average hence resulting in more responses that mistake a set average as the set member. The second experiment investigated whether the default processing bias could be altered when a global or local processing orientation is encouraged through priming. We tested East Asian participants on the same task as Experiment 1, but before the task the participants were either primed with a global or local processing orientation. Experiment 2 tested the hypothesis that the cultural default mode of processing is to some extent malleable in practice, within a processing context. The relative emphasis on global or local information, or the relative focus on the overall set average or set member should also depend on the external demand for a specific processing orientation.

Experiment 1

We tested a group of East Asian (Chinese) and a group of Western (British) participants on de Fockert and Wolfenstein's (2009) ensemble coding task. Given that East Asians are more inclined to utilize global processing, we hypothesized that such processing orientation could lead to more frequent misidentification of a set average as a set member for East Asian than for Western participants in this task.

Methods

Participants. Forty-seven East Asian students from Renmin University of China (20.17±2.24 years old) and 48 Caucasian students from Bournemouth University of Britain (19.47±1.04 years old) participated in this experiment. All experiments in this study were approved by the local Institutional Review Board of both institutions. Written consent was obtained from all participants.

Materials. Stimuli were 120 Asian and 120 White faces collected from the internet. The background and the external features of each face were masked out in Photoshop. Both the Asian and White stimuli consisted of an equal number of male and female faces. The 60 faces within each ethnicity and gender were then randomly assigned to 15 sets of 4 faces. Following de Fockert and Wolfenstein (2009), we created a total of 60 morphed faces. Each morphed face was created from the four original faces in each set using the Abrosoft FantaMorph 5.0. To create a morphed face in the software, a total of 120 markers that define the key facial features (i.e., the contour of the eyes, the eyebrows, the nose, the mouth, and the facial outline) of each face were manually

selected. After this step, the software would linearly calculate the mean RGB values of the corresponding-points on the four faces into a single morph.

Each face image (170 × 220 px) subtended a visual angle of approximately 4.8° × 7.4° at a viewing distance of around 60cm, and was presented against a black background. An example of a face set is given in Figure 1.
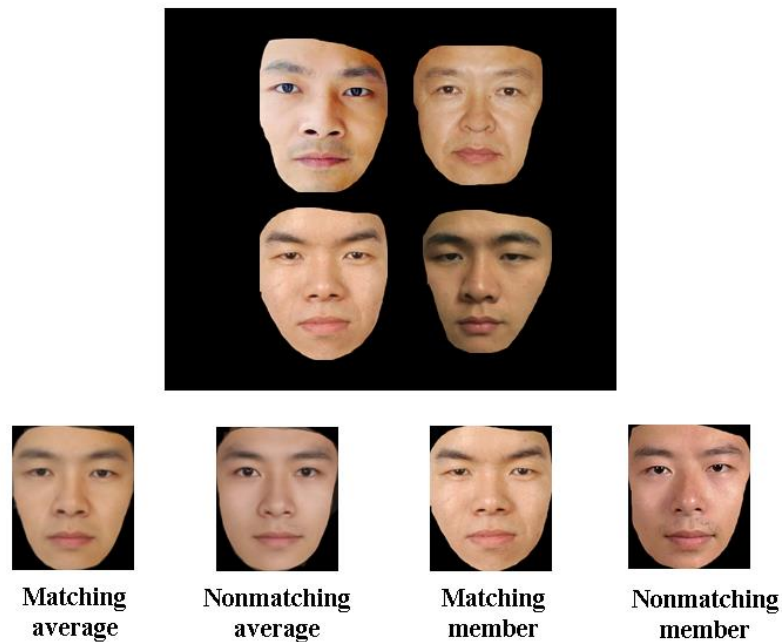

Figure 1. Example of a face set and types of test faces used in Experiment 1. The image at the top shows an example of face set. The images at the bottom illustrate the four types of test image: a morphed average identity of the preceding set (matching member), a morphed average identity of another face set (nonmatching member), a member of preceding set (matching member) and a member of another face set (nonmatching member). (Note: The figure is created for illustration purpose only.)

Procedure. The procedure was identical to that of de Fockert and Wolfenstein (2009). Each trial began with a 500ms fixation cross, followed by a set of 4 faces presented for 2000ms. This was then replaced by a single test face, which remained on the screen until a response was made. Participants were asked to judge whether or not the test face was presented in the preceding set. They pressed "f" key with the left index finger for "present", and "j" key with the right index finger for "absent". As depicted in Figure 1, the test face consisted of four types: a morphed average of the four faces in the preceding set (matching average), a morphed average of another set of the same ethnicity and gender (nonmatching average), a member of the preceding set

(matching member), or a member of another set of the same ethnicity and gender (nonmatching member). Following de Fockert and Wolfenstein (2009) and de Fockert and Gautrey (2013) the nonmatching average and nonmatching member in each trial were selected from different face sets.

The experiment was blocked by face ethnicity. The order of Asian and White face blocks was counterbalanced across participants. The two face genders and the four test-image types were randomized within each block. The position of the four images in each set was also randomized from one trial to another. Since there were 15 sets for each face gender and test-image type, it amounted to 120 trials (15 set × 2 face gender × 4 test image type) types. The same 120 trials were repeated once after their completion. This resulted in a total of 240 trials per block. Participants were given 10s rest after every 60 trials. A practice block with 15 trials was conducted to familiarize participants with experiment procedure. The procedure was implemented with and run by E-prime 2.0 and presented on a 1024 × 768 px screen.

Design. We employed a mixed-factor design. Participant Ethnicity (East Asian vs. Western) was a between-participants variable, whereas Test-image Type (Set average vs. Set members) and Face Ethnicity (Asian vs. White) were within-participants variables. Following Rhodes et al. (2015), we employed "endorsement score" as the dependent variable. This score was calculated by subtracting the "present" responses in nonmatching conditions from matching conditions for both set average and set members. It served to assess unbiased recognition for matching member and matching average, without directly comparing the performance on averages and members that might differ in low-level image properties (smoother skin texture on averages than members). The value of the "endorsement score" measures the strength of ensemble coding of multiple face identities when the test face is a set average. It indicates the level of accuracy when the test face is a set member. A positive endorsement score indicates a judgment based on the identity information in the study set, whereas zero indicates chance performance. A negative score suggests below-chance performance. Descriptive data concerning the mean "present" responses under each condition as well as the endorsement scores was displayed in Table 1.

Table 1. Descriptive data concerning the mean proportion of "present" response and endorsement scores under each condition. *M(SD)*

		Proportion of "present" responses				Endorsement scores	
		Matching average	Nonmatching average	Matching member	Nonmatching member	Set average	Set member
East Asian participants	Asian face	.66(.14)	.26(.12)	.63(.14)	.14(.08)	.40(.09)	.48(.12)
	White face	.74(.17)	.20(.13)	.62(.15)	.12(.10)	.54(.15)	.50(.13)
Western participants	Asian face	.63(.16)	.25(.12)	.58(.12)	.18(.11)	.37(.16)	.40(.14)
	White face	.63(.20)	.24(.17)	.59(.14)	.20(.15)	.39(.21)	.39(.18)

Results and Discussion

Results are shown in Figure 2. A three-way repeated-measures Analysis of Variance (ANOVA) found a significant effect of Participant Ethnicity, $F (1, 93) = 13.31$, $p < .001$, $\eta^2 = .13$, where East Asian participants showed higher endorsement scores ($M = .48$) than Western participants ($M = .39$). There was also a significant effect of Face Ethnicity, $F (1, 93) = 11.73$, $p = .001$, $\eta^2 = .11$, where Asian faces created lower endorsement scores ($M = .41$) than White faces ($M = .46$). The main effect of Test-image Type was not significant, $F (1, 93) = 3.04$, $p = .08$, $\eta^2 = .03$. These main effects were qualified by two-way and three-way interactions: Face Ethnicity × Test-image Type ($F (1, 93) = 18.38$, $p < .001$, $\eta^2 = .17$; Participant Ethnicity × Face Ethnicity ($F (1, 93) = 10.42$, $p = .002$, $\eta^2 = .10$; and Face Ethnicity × Test-image Type × Participant Ethnicity, $F (1, 93) = 6.23$, $p = .01$, $\eta^2 = .06$.

Post-hoc LSD tests indicated that: (1) East Asian participants showed higher endorsement scores than their Western counterparts when the test face was own-ethnicity member ($F (1, 93) = 9.27$, $p = .003$), or other-ethnicity member ($F (1, 93) = 12.97$, $p = .001$) or when the test face was an other-ethnicity average ($F (1, 93) = 15.18$, $p < .001$), although not so for an own-ethnicity average ($F (1, 93) = .76$, $p = .39$): (2) East Asian participants showed lower endorsement scores when the test face was own-ethnicity average than other-ethnicity average ($F (1, 93) = 35.67$, $p < .001$), but equal endorsement scores for own- and other-ethnicity members ($F (1, 93) = 1.18$, $p = .279$); Western participants showed no difference between their endorsement scores for own-

and other-ethnicity faces, whether the test face was a set average (F (1, 93) = .61, p = .44) or a set member (F (1, 93) = .52, p = .47).
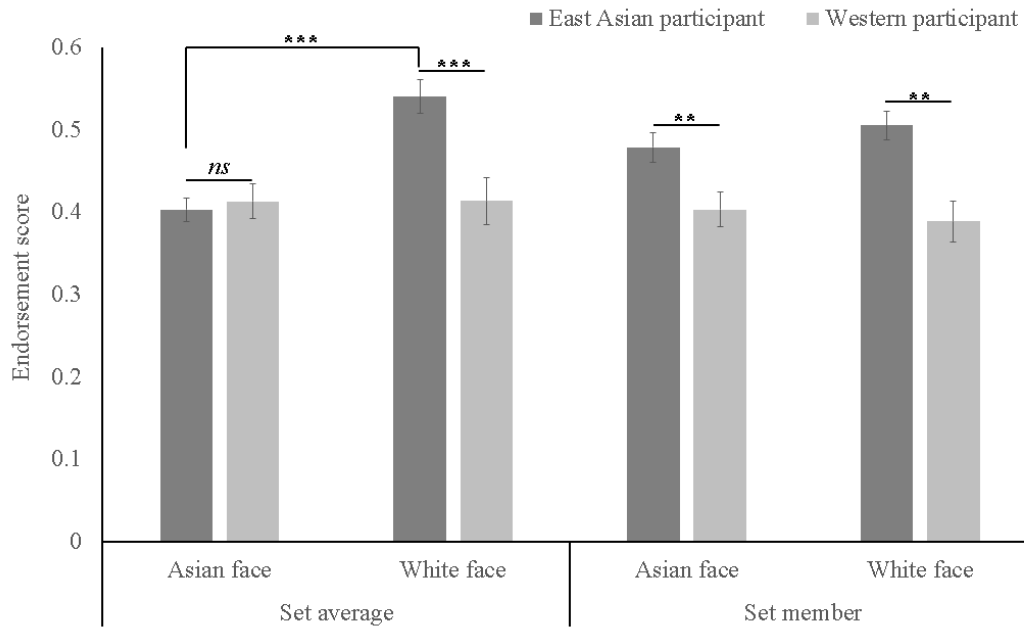

Figure 2. Results of Experiment 1 as a function of Participant Ethnicity, Face Ethnicity, and Test-image Type. Error bars indicate one standard error about the mean.

The key result thus shows that East Asian participants produced higher endorsement scores for set average than their Western counterparts, which was mainly due to their higher scores for other-ethnicity than for own-ethnicity faces. In other words, there was an asymmetric effect in ensemble coding of own- versus other-ethnicity face: while Western participants showed comparable set average endorsement scores for own- and other-ethnicity faces, East Asians displayed an other-ethnicity bias in face ensemble coding. This is largely consistent with the hypothesis that East Asians may be more prone to ensemble coding relative to Westerners. The stronger tendency to employ ensemble coding could reflect a default mode of East Asian observers. It may be due to a stronger bias for global than local information. However, how rigid is this bias? Given that local information could also be important from time to time in daily life, can the default mode be modulated to adapt to a context-dependent demand? Experiment 2 investigated this question.

Experiment 2

Experiment 1 showed cultural influence on a stable processing mode for ensemble coding. However, although East Asian observers may tend to focus more on holistic information than observers from a Western culture, it would also be necessary for them to attend to local information from time to time. How malleable is the default processing mode? To examine the hypothesis that the default mode could be adjusted by a momentary shift of processing orientation, we employed a Navon letter task (Navon, 1977) to prime East Asian (Chinese) participants' global or local processing orientation before having them perform the same ensemble coding task in this experiment. A global processing orientation should encourage individuals to focus more broadly on the ensemble properties of a visual field, while a local processing orientation should direct individuals to relatively local details such as the constituents in a stimulus set. Since East Asian participants are already prone to holistic processing, we hypothesized that priming global processing orientation should have a relatively small effect on further bias for set average, whereas local processing orientation should reduce the proneness to ensemble coding.

Methods

Participants. Sixty college students from Renmin University of China were randomly assigned to one of the two priming conditions. Both groups had 30 participants. The age was 19.9 ± 1.3 years old in the global priming condition, and 20.8 ± 2.54 years old in the local priming condition. Some of the participants (N=16) in this experiment also participated in Experiment 1, and half of them was in the global priming condition and the other half was in the local priming condition. Written consent was obtained from all participants.

Stimuli. Because those who participated in Experiment 1 were also allowed to participate in this experiment, we decided to use a different set of Asian faces to minimise the effects of practice and familiarity. A total of 56 Asian faces (half males and half females) were selected from the Chicago Face Database (Ma et al., 2015). External features (hair, clothing, etc.) were occluded via an oval mask. Every 4 faces of the same gender were randomly assigned to a set, which resulted in 14 sets of faces. We created 14 morphed faces, averaging the four faces in each set. Each face image was presented against a black background, measuring at 170 × 220 px, or 4.8° × 7.4° of visual angle at a distance of around 60cm.

Stimuli in the Navon letter task consisted of alphabetical letters. Each global letter consisted

of small local letters. See the second frame in Figure 3 for an example. Here the global letter is 'H', which consists of small local letters 'K'. The Navon letter task in the global priming condition contained global letters 'A' and 'H'. They consisted of local letters 'K', 'T', 'L' and 'E', The task in the local priming condition consisted of local letters A and H, with each forming global letters 'F', 'K', 'T' and 'E'. All stimuli were presented against a black background, with each local letter subtended 0.7° × 0.9° and each global letter occupied around 3 × 4.8° of visual angle.
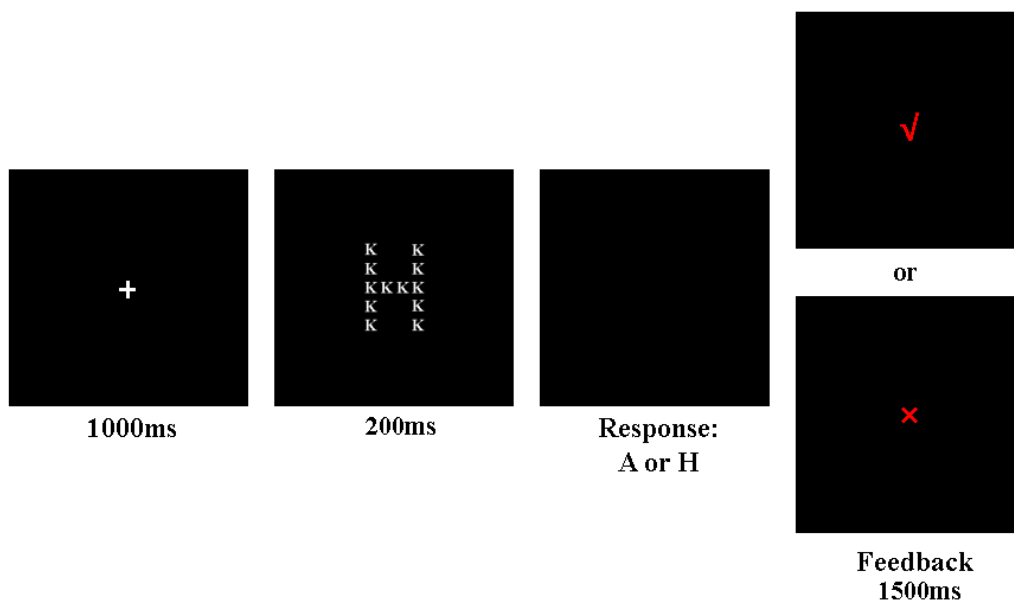

Figure 3. Example of a priming trial. Following a 1000ms fixation, a Navon letter array was presented (here a global H composed of local Ks) for 200ms. In this example, participants judged whether the global letter was A or H. Once a response was made, a 1500-ms feedback display was given.

Procedure. The procedure was identical to that of Experiment 1, except that a priming task was added before the face recognition task. As shown in Figure 4, the priming task began with a 1000ms fixation, followed by a Navon letter stimulus at the center of the screen for 200ms. Participants were required to judge whether they saw an A or H. An "A" key press indicated the presentation of letter A, while an "H" key press indicated the presentation of letter H. Participants in the global-priming condition were required to respond to the big letter made of the small ones, while participants in the local-priming condition were told to respond to the small letter. The display was blank until a response was made. Feedback was given following the response.

The priming task had a total of 32 trials (8 letters × 4 repetitions). It was then followed by the ensemble coding task. To ensure a sustainable effect of priming, participants repeated the same priming task but with only 16 trials (8 letters × 2 repetitions) after they finished each block. In total, this experiment contained 8 blocks[1] of 56 trials (14 face sets × 4 types of test image). The experimental procedure was implemented with and run by E-prime 2.0 and presented on a 1024 × 768 px screen.

Design. We employed a mixed-factor design. Priming Condition (global vs. local) was a between-participants variable, whereas Test-image Type (Set average vs. Set member) was a within-participants variable. Endorsement score was the dependent variable.

Results and Discussion

The overall accuracy of Navon letter task was higher than 95% for each participant. There was no difference between the accuracies for the global letter condition ($M = 98.6\%$) and the local letter condition ($M = 98.0\%$), $t(58) = 1.48$, $p = .15$. Results of the recognition task are shown in Figure 4. ANOVA showed a significant main effect of Test-image Type, $F(1, 58) = 16.01$, $p < .001$, $\eta^2 = .22$), where participants produced more "present" responses to set members ($M = .36$) than to set averages ($M = .28$). The main effect of Priming Condition was also significant, $F(1, 58) = 21.76$, $p < .001$, $\eta^2 = .27$, where global priming ($M = .39$) created a higher endorsement score for set averages than local priming ($M = .25$). These main effects, however, were qualified by a significant two-way interaction, $F(1, 58) = 30.89$, $p < .001$, $\eta^2 = .35$.
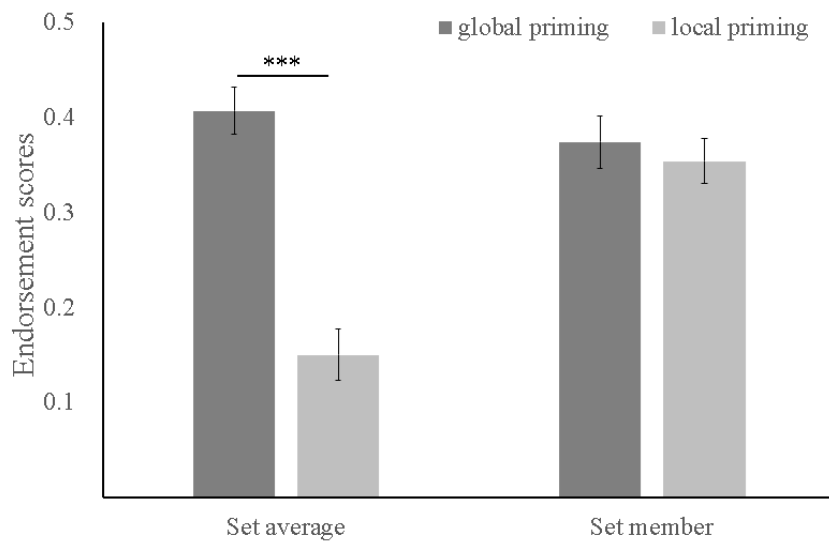
Figure 4.   Results of Experiment 2 as a function of Priming Condition and Test-image Type. Error bars represent one standard error about the mean.

Post-hoc LSD analyses showed that participants primed with global processing orientation (M = .41) were more likely to endorse set average as a group member of the preceding set than those primed with local processing orientation (M = .15), $p < .001$. In contrast, no difference was found between results of global and local priming condition (M's = .37 and .35, respectively) for the set members, $p = .59$. Additionally, set member was more strongly endorsed than set average after local priming, $p < .001$, while no such difference was found between the results of the set average and set members after global priming, $p = .28$.

As predicted, the results of this experiment demonstrate that the default bias for ensemble coding in a cultural group could be temporarily altered via priming. After being primed with a local processing orientation, East Asian participants significantly reduced their inclination for ensemble coding.

General Discussion

We conducted two experiments to explore how culture and priming for local or global processing orientation affect ensemble coding. Experiment 1 compared an East Asian (Chinese) group with a Western (British) group on their ensemble coding of multiple face identities. Our results replicated the finding of de Fockert and Wolfenstein (2009), showing that participants in

both cultural groups often treated the set average as a set member of a preceding face set. More importantly, however, we found the East Asian participants showed an overall stronger tendency for ensemble coding by choosing more set averages as a set member than Western participants. East Asian participants tended to choose a set average of other-ethnicity faces more often relative to a set average of own-ethnicity faces. The reverse was not true among Western participants. The result confirms our hypothesis that the culture can influence the propensity for ensemble coding.

Experiment 2 primed East Asian (Chinese) participants with either global or local processing orientation before performing the same task as Experiment 1. Results showed that priming of local processing orientation reduced endorsement for set averages relative to priming of global processing orientation, thus confirming the hypothesis that the default cultural mode could be temporarily shifted to a degree.

The results of Experiment 1 demonstrated that relative to the Caucasian participants, Asian participants had a stronger tendency to average White faces. This kind of culture difference in ensemble perception is consistent with Im et al. (2017), who found that Korean participants performed better than American in extracting the average emotion of a facial crowd. However, as Im et al. acknowledge themselves, their outcomes could also be explained by a motivational difference between the two cultural groups, as East Asians usually focus more on avoidance-orientated information while Westerners focus more on approach-orientated information. Our study should have ruled out this explanation, because the paradigm provided more direct evidence for the cultural effect on ensemble coding. Instead, the results may be better explained by a greater focus of East Asian participants on global information.

The results of Experiment 1 appeared to contradict with the previous finding that stronger holistic processing is tuned to own-ethnicity faces (e.g., Michel et al., 2006). However, the literature on this has created mixed results. For example, some studies have shown that East Asians displayed comparable holistic processing for both own-ethnicity and other-ethnicity faces, while Westerners displayed stronger global processing of own-ethnicity faces (Tanaka et al., 2004; Michel et al., 2006; Crookes et al., 2013). To note here, previous studies have mainly investigated the interaction between holistic processing and own-/other-ethnicity in individual face perception. The relationship between global processing and ensemble perception (multiple faces perception) of own- versus other-ethnicity faces is also not always clear-cut. In Im et al. (2017), East Asians

showed comparable accuracy in averaging own-ethnicity and other-ethnicity faces in Experiment 1, but produced better averaging performance for other-ethnicity than for own-ethnicity faces in Experiment 2, where response speed was emphasized. In the present study, we also found that East Asian participants showed stronger sensitivity to average face of other-ethnicity than own-ethnicity faces, which was the major reason for the stronger ensemble perception of East Asians relative Westerners. Im et al found North American showed better performance in averaging of own-ethnicity faces than other-ethnicity faces in both their Experiments 1 and 2, while in our study we found Westerners showed equal propensity for average face of own-ethnicity and other-ethnicity faces. Hence, whether global (versus analytic) processing influence ensemble coding of own- and other- ethnicity facial crowds will require further investigation.

Global processing orientation requires spreading the attentional resources over a large field, whereas local processing orientation requires attentional resources to be more narrowly focused on an object or a region, which enhances the precision of a smaller visual field at the expense of other objects and field in vision (Eriksen & St. James, 1986; Müller, et al., 2003; Srikanthan et al., 2018). Cultural differences in breadth of attention have been investigated in prior research (Kitayama & Murata, 2013; Park & Kitayama, 2014), which has shown that Westerners tend to be more narrowly focused while Eastern Asians tend to be more broadly and holistically focused on visual stimuli. We suggest that the global/local attention bias is linked to global/local processing orientation for ensemble coding, which may account for the cultural difference found in this study.

This study also demonstrates that the cultural influence on ensemble coding is not rigid. Experiment 2 showed that the cultural bias could be modulated by temporarily adapting to a more global or local processing orientation. The momentary alteration between global and local processing orientation may reflect the effects of variable attentional scope (Sun et al., 2016). A stimulus set is usually treated as a whole rather than individual items when the set average is estimated. Chong and colleagues (2008) describe this as a distributed attention mode. It is plausible that activation of a local processing orientation, which deviates from the default distributed attention mode, could impair ensemble coding. This may explain the priming effect found in Experiment 2, where attention to local Navon letters could have led to reduced responses for set averages. The priming of global processing orientation, however, did not create more

responses for set averages. This could be due to an already fairly strong global processing bias among the Asian participants.

Similar to studies that demonstrate priming effects on interdependent and independent self, in which temporary access to other cultural frameworks can alter a default response to one's own and friend's faces (Sui & Han, 2007; Sui et al., 2013), our study demonstrate that there is a degree of flexibility when individuals in a culture respond to global or local information based on a specific demand at hand. A shortcoming of this study was it did not include a baseline condition. This makes it difficult to ascertain whether our effect was due to facilitation from global priming or inhibition from local priming. Based on a causal comparison of our results in the two experiments, it seems possible that local priming inhibited averaging, because the endorsement scores by the Chinese participants for Asian face set averages were very similar either without global priming ($M = .40$) or after global priming ($M = .39$). However, only a rigorous test with a proper baseline can really answer the question. Another shortcoming was that we did not include Western group in Experiment 2. However, based on previous studies that used a priming task on both East Asian and Western participants to elicit either a collectivist or an individualistic orientation (Kühnen & Oyserman, 2002; Oyserman & Lee, 2008; Sui et al., 2013), it is quite possible that a similar priming effect in our experiment is also found in a Western sample.

Another potential issue of our study is that our findings could also be explained in terms of object and feature interactions in working memory. Prior research has demonstrated that the individual features of presented faces could form a set of combined representations that have to be distinguished from an ensemble (Reinitz et al., 1992). Based on this, it may be argued that the East Asians were influenced by ensemble features to different degrees depending on the stimulus type (Asian vs. White faces) can thus be in line with an idea of such false memories. For example, if East Asians remember White faces less holistically, then it is more likely that they mix up the features to form a false memory. Similarly, primed by a local attention task, participants can encode each face more separately from others, thus providing less "swaps" between different faces (Treisman, 2006). However, this interpretation does not easily apply to the results of our Western participants, who showed comparable tendency to endorse Asian and White faces. At this point, it would be difficult to infer from the current results alone about whether the differences between the results of our two groups could be explained by feature binding in working memory. A test of this

alternative explanation can only be achieved by a systematic manipulation of the individual features in the face stimuli, which is worthy of pursuing in future research.

Since both East Asian and Western participants are able to rapidly extract an average of multiple faces (de Fockert & Wolfenstein, 2009; Li et al., 2016), ensemble coding is likely a strategy for visual information processing across cultures. Consistent with prior research, both East Asian and Western participants in our study showed a tendency to average face identity in perceiving facial crowds. The results support the suggestion that ensemble representation is a universal information processing mechanism with evolutionary significance (Alverez, 2011). However, along with Im et al. (2017), our study has provided the added evidence that culture can modulate the focus on ensemble representations. Moreover, our study has also shown that the default mode of processing orientation in a culture can be affected by specific task demands hence is malleable in practice. Future research will need to find out the trajectory of the cultural difference and the mechanisms of attentional modulation between global and local processing orientations.

Notes

1: The kind of repeated presentation of the same stimuli was adopted in previous studies (de Fockert & Wolfenstein, 2009; Haberman et al.,2015). See, for example, Study 1 of Haberman et al. (2015), where "observers saw a total of 20 unique face sets and 20 unique gabor sets *8* times each". Haberman et al. (2015) computed the reliability of the ensemble coding task and found that the Cronbach's alpha was .84 in their ensemble face task, suggesting that averaging of multiple face identities was highly consistent. However, according to the results of the ANOVA based on the first block in Experiment 2, we found that though the result pattern did keep similar, the effect size of the Priming Condition × Test-image Type interaction based on the first block (F (1, 58) = 8.55, p = .005, η^2 = .13) was smaller than that based on the all trials (η^2 = .35) as shown in the main text.

References

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, *15*(3), 122-131.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, *12*(2), 157-162.

Avidan, G., Tanzer, M., & Behrmann, M. (2011). Impaired holistic processing in congenital prosopagnosia. *Neuropsychologia*, *49*(9), 2541-2552.

Chen, B., & Zhou, G. (2018). Attentional modulation of hierarchical ensemble coding for the identities of moving faces. *Journal of Experiment Psychology: Human Perception and Performance, 44*(10), 1542-1556.

Chong, S. C., Joo, S. J., Emmmanouil, T. A., & Treisman, A. (2008). Statistical processing: Not so implausible after all. *Perception & Psychophysics*, *70*(7), 1327-1334.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*(4), 393-404.

Chong, S. C., & Treisman, A. (2005a). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics*, *67*(1), 1-13.

Chong, S. C., & Treisman, A. (2005). Statistical processing: Computing the average size in perceptual groups. *Vision Research*, *45*(7), 891-900.

Crookes, K., Favelle, S., & Hayward, W. (2013). Holistic processing for other-race faces in Chinese participants occurs for upright but not inverted faces. *Frontiers in Psychology*, *4*, 29.

de Fockert, J. W., & Gautrey, B. (2013). Greater visual averaging of face identity for own-gender faces. *Psychonomic Bulletin & Review*, *20*(3), 468-473.

de Fockert, J., & Wolfenstein, C. (2009). Rapid extraction of mean identity from sets of faces.

Quarterly Journal of Experimental Psychology, *62*(9), 1716-1722.

Elliot, A. J., Chirkov, V. I., Kim, Y., & Sheldon, K. M. (2001). A cross-cultural analysis of avoidance (relative to approach) personal goals. *Psychological Science*, *12*(6), 505-510.

Eriksen, C. W., & James, J D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Percept & Psychophysics, 40*(4), 225-240.

Haberman, J., Brady, T. F., & Alvarez, G. A. (2015). Individual differences in ensemble perception reveal multiple, independent levels of ensemble representation. *Journal of Experimental Psychology: General*, *144*(2), 432-446.

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology*, *17*(17), R751-R753.

Haberman, J., & Whitney, D. (2009). Seeing the mean: ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 718-734.

Hamamura, T., Meijer, Z., Heine, S. J., Kamaya, K., & Hori, I. (2009). Approach—Avoidance motivation and information processing: A cross-cultural analysis. *Personality and Social Psychology Bulletin*, *35*(4), 454-462.

Han, S., Northoff, G., Vogeley, K., Wexler, B. E., Kitayama, S., & Varnum, M. E. (2013). A cultural neuroscience approach to the biosocial nature of the human brain. *Annual Review of Psychology*, *64*, 335-359.

Im, H. Y., Chong, S. C., Sun, J., Steiner, T. G., Albohn, D. N., Adams, R. B., & Kveraga, K. (2017). Cross-cultural and hemispheric laterality effects on the ensemble coding of emotion in facial crowds. *Culture and Brain*, *5*(2), 125-152.

Ji, L., Chen, W., & Fu, X. (2014, June 22-27). Different roles of foveal and extrafoveal vision in ensemble representation for facial expressions. In 11th *International Conference on Engineering Psychology and Cognitive Ergonomics* (pp. 164–173). Springer, Cham.

Ji, L. J., Peng, K., & Nisbett, R. E. (2000). Culture, control, and perception of relationships in the environment. *Journal of Personality and Social Psychology*, *78*(5), 943-955.

Kramer, R. S., Ritchie, K. L., & Burton, A. M. (2015). Viewers extract the mean from images of the same person: A route to face learning. *Journal of Vision*, *15*(4):1, 1-9.

Kitayama, S., & Murata, A. (2013). Culture modulates perceptual attention: An event-related potential study. *Social Cognition*, *31*(6), 758-769.

Kühnen, U., & Oyserman, D. (2002). Thinking about the self influences thinking in general: cognitive consequences of salient self-concept. *Journal of Experimental Social Psychology, 38*(5), 492–499.

Lao, J., Vizioli, L., & Caldara, R. (2013). Culture modulates the temporal dynamics of global/local processing. *Culture and Brain*, *1*(2-4), 158-174.

Lewis, R. S., Goto, S. G., & Kong, L. L. (2008). Culture and context: East Asian American and European American differences in P3 event-related potentials and self-construal. *Personality and Social Psychology Bulletin*, *34*(5), 623-634.

Li, H., Ji, L., Tong, K., Ren, N., Chen, W., Liu, C. H., & Fu, X. (2016). Processing of individual items during ensemble coding of facial expressions. *Frontiers in psychology*, *7*, 1332.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122-1135.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*(2), 224-253.

Masuda, T., Ellsworth, P. C., Mesquita, B., Leu, J., Tanida, S., & Van de Veerdonk, E. (2008). Placing the face in context: Cultural differences in the perception of facial emotion. *Journal of Personality and Social Psychology, 94*(3), 94365–94381.

Masuda, T., & Nisbett, R. E. (2001). Attending holistically versus analytically: comparing the context sensitivity of Japanese and Americans. *Journal of Personality and Social Psychology, 81*(5), 922-934.

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

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

Miyamoto, Y. (2013). Chapter Three – Culture and Analytic Versus Holistic Cognition: Toward Multilevel Analyses of Cultural Influences. *Advances in Experimental Social Psychology*, 47, 131–188.

Müller, N. G., Bartelt, O. A., Donner, T. H., Villringer, A., & Brandt, S. A. (2003). A physiological correlate of the "zoom lens" of visual attention. *Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 23*(9), 3561-3565.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*(3), 353-383.

Neumann, M. F., Schweinberger, S. R., & Burton, A. M. (2013). Viewers extract mean and individual identity from sets of famous faces. *Cognition*, *128*(1), 56-63.

Neumann, M. F., Ng, R., Rhodes, G., & Palermo, R. (2018). Ensemble coding of face identity is not independent of the coding of individual identity. *The Quarterly Journal of Experimental Psychology, 71*(6), 1357–1366.

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

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review, 108*(2), 291-310.

Oyserman, D., & Lee, S. W. S. (2008). Does culture influence what and how we think? Effects of priming individualism and collectivism. *Psychological Bulletin, 134*(2), 311–342.

Park, J., & Kitayama, S. (2014). Interdependent selves show face-induced facilitation of error processing: cultural neuroscience of self-threat. *Social Cognitive and Affective Neuroscience*, *9*(2), 201-208.

Pavlovskaya, M. , Soroker, N. , Bonneh, Y. S. , & Hochstein, S. . (2015). Computing an average when part of the population is not perceived. *Journal of Cognitive Neuroscience, 27*(7), 1397-1411.

Peng, S., Zhang, L., Xu, R., Liu, C. H., Chen, W. & Hu, P. (2019). Self-Construal Priming Modulates Ensemble Perception of Multiple-Face Identities. *Frontiers in Psychology*, *10*, 1096.

Reinitz, M.T., Lammers, W.J. & Cochran, a. P. (1992) Memory-conjunction errors: Miscombination of stored stimulus features can produce illusions of memory. *Memory & Cognition, 20*(1), 1–11.

Robson, M. K., Palermo, R., Jeffery, L., & Neumann, M. F. (2018). Ensemble coding of face identity is present but weaker in congenital prosopagnosia. *Neuropsychologia, 111*, 377–386.

Rhodes, G., Neumann, M. F., Ewing, L., & Palermo, R. (2015). Reduced set averaging of face identity in children and adolescents with autism. *The Quarterly Journal of Experimental Psychology*, *68*(7), 1391-1403.

Rhodes, G., Neumann, M., Ewing, L., Bank, S., Read, A., Engfors, L. M., ... & Palermo, R. (2017). Ensemble coding of faces occurs in children and develops dissociably from coding of individual faces. *Developmental Science*, *21*(2), e12540.

Srikanthan, D., Sama, M., Sun, S., Nestor, A., & Cant, J. (2018). Reciprocal Interference between Global and Local Processing in Ensemble Perception. *Journal of Vision*, *18*(10), 76-76.s

Sui, J., & Han, S. (2007). Self-construal priming modulates neural substrates of self-awareness. *Psychological Science*, *18*(10), 861-866.

Sui, J., Hong, Y. Y., Liu, C. H., Humphreys, G. W., & Han, S. (2012). Dynamic cultural modulation of neural responses to one's own and friend's faces. *Social Cognitive and Affective Neuroscience*, *8*(3), 326-332.

Sui, J., Liu, C. H., & Han, S. (2009). Cultural difference in neural mechanisms of self-recognition. *Social Neuroscience*, *4*(5), 402-411.

Sun, S. Z., Cant, J. S., & Ferber, S. (2016). A global attentional scope setting prioritizes faces for conscious detection. *Journal of Vision*, *16*(6):9, 19-13.

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

Treisman, A. (2006). How the deployment of attention determines what we see. *Visual Cognition, 14*(4-8), 411–443.

Triandis, H. C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review*, *96*(3), 506-520.

Whitney, D., & Leib, A. Y. (2018). Ensemble perception. *Annual Review of Psychology*, *69*, 105-129.