



Processing of facial expressions of same-race and other-race faces: distinct and shared neural underpinnings

Xuena Wang  and Shihui Han 

School of Psychological and Cognitive Sciences, PKU-IDG/McGovern Institute for Brain Research, Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing 100080, China

Correspondence should be addressed to Shihui Han, School of Psychological and Cognitive Sciences, PKU-IDG/McGovern Institute for Brain Research, Peking University, No. 52, Haidian Street, Beijing 100080, China. E-mail: shan@pku.edu.cn.

Abstract

People understand others' emotions quickly from their facial expressions. However, facial expressions of ingroup and outgroup members may signal different social information and thus be mediated by distinct neural activities. We investigated whether there are distinct neuronal responses to fearful and happy expressions of same-race (SR) and other-race (OR) faces. We recorded electroencephalogram from Chinese adults when viewing an adaptor face (with fearful/neutral expressions in Experiment 1 but happy/neutral expressions in Experiment 2) and a target face (with fearful expressions in Experiment 1 but happy expressions in Experiment 2) presented in rapid succession. We found that both fearful and happy (*vs* neutral) adaptor faces increased the amplitude of a frontocentral positivity (P2). However, a fearful but not happy (*vs* neutral) adaptor face decreased the P2 amplitudes to target faces, and this repetition suppression (RS) effect occurred when adaptor and target faces were of the same race but not when of different races. RS was observed on two late parietal/central positive activities to fearful/happy target faces, which, however, occurred regardless of whether adaptor and target faces were of the same or different races. Our findings suggest that early affective processing of fearful expressions may engage distinct neural activities for SR and OR faces.

Key words: EEG; fear; happy; race; repetition suppression

Introduction

It has been realized since Darwin (1872) that facial expressions evolved to deliver critical social information that relates to human survival (Shariff and Tracy, 2011). The human brain has developed specific psychological and neural mechanisms underlying the adaptive function of expression processing. Research measuring behavioral performance (see Calvo and Nummenmaa, 2016 for review) and facial electromyography (Kret *et al.*, 2013) has shown evidence that exposure to both positive (e.g. happy) and negative (e.g. fear, angry) facial expressions

generates immediate and unintentional affective responses in humans. Functional magnetic resonance imaging (fMRI) studies further discovered multiple neural circuits involved in automatic processing of facial expressions. For example, fearful faces most strongly activate the amygdala, and happy faces most strongly activate the superior temporal gyrus and rostral anterior cingulate cortex (see Vytal and Hamann, 2010 for a meta-analysis).

Most of the previous neuroimaging studies focused on what and how perceptual features contribute to recognition of facial

Received: 4 August 2020; Revised: 23 December 2020; Accepted: 24 February 2021

© The Author(s) 2021. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

expressions and how facial expressions in turn generate emotional responses (Calvo and Nummenmaa, 2016). To date, much less is known about the relationship between discrepant adaptive functions of various expressions and neural mechanisms underlying the processing of these facial expressions in complicated social contexts. For example, we have known little about whether the same facial expression (e.g. fear) displayed by individuals from different social groups (e.g. ingroup or outgroup) is encoded by the same or distinct neural activities. It is important to address this issue because the same expression displayed in an ingroup member or an outgroup member may deliver different nonverbal social signals about intended actions pertaining to an observer's safety and survival.

Previous research has shown that fearful compared to happy expressions are more easily paired with aversive stimuli (Öhman and Mineka, 2001). Even infants can take parents' fearful expressions as signals to guide their own decisions regarding whether to cross a visual cliff (Sorce et al., 1985), suggesting that fearful expressions can alert possible threats to observers. However, fearful expressions of ingroup and outgroup members may have different social meanings. For example, in a context of intergroup conflict, observing an ingroup member's fearful expression may provide a signal of threats to all ingroup members. The same fearful expression shown by an outgroup member, however, may implicate the weakness of the outgroup relative to the ingroup and hence signal safety for the ingroup and oneself (Weisbuch and Ambady, 2008). From an adaptive perspective, perceived threats or weakness signaled by fearful expressions have different meanings for an observer's safety and survival and thus require different reactions. It is thus likely that fearful expressions of ingroup and outgroup members may be encoded by distinct neural activities at some stages of processes of facial expressions. A neural strategy like this ought to benefit an observer by taking different and quick reactions to fearful expressions perceived from ingroup and outgroup members.

Unlike fearful expressions, happy expressions communicate a lack of threats (Preuschoft and van Hooff, 1997; Shariff and Tracy, 2011). Behavior tendency to mimic smiling can overrule group boundaries (Bourgeois and Hess, 2008; van der Valk et al., 2011). If happy expressions shown by ingroup and outgroup members do not signal any threats (Weisbuch and Ambady, 2008), the brain may not develop distinct neural activities to encode happy expressions of ingroup and outgroup members. In other words, the processing of happy expressions of ingroup and outgroup members may engage shared neural underpinnings.

Consistent with these hypotheses, a review of multiple studies of different cultural samples has shown that the mean recognition agreement scores across expressions are much higher for happy than for fearful expressions (Nelson and Russell, 2013). Moreover, individuals typically respond to happy expressions more accurately and faster compared to all the other expressions, whereas they show the poorest accuracy and longest latencies for fearful faces (Calvo and Nummenmaa, 2016). These behavioral findings imply homogeneous processing of happy expressions across cultures and a higher congruency in social information delivered by happy compared to fearful expressions. More relevant to our hypotheses, previous brain imaging studies have examined neural responses to expressions of same-race (SR) and other-race (OR) faces. The categorization of race has been supposed to be a byproduct of evolved mechanisms underlying identification of others' coalitional status (e.g. perceive SR faces as ingroup and OR faces as outgroup) (Kurzban et al., 2001; Cosmides et al., 2003). Social categorization of faces

by race occurs spontaneously and early in the ventral temporal cortex and influences social emotion (e.g. empathy) and altruistic behavior (Han, 2018; Zhou et al., 2020). fMRI studies found that Japanese and white participants showed greater amygdala responses to fearful expressions of SR faces compared to those of OR faces, whereas happy expressions of SR and OR faces did not differentially activate the brain regions typically engaged in the processing of happy expressions such as the superior temporal gyrus and rostral anterior cingulate cortex (Chiao et al., 2008; Iidaka et al., 2008). These neuroimaging results, although providing evidence for discrepant neural responses to fearful expressions of SR and OR faces, are unable to verify whether distinct neural activities are engaged in the processing of fearful expressions of SR and OR faces. This is because the observed greater amygdala responses to fearful expressions of SR (vs OR) faces might reflect modulations of the same neural activity by perceived race.

Our recent ERP study developed a paradigm that allowed examination of whether distinct neuronal populations are engaged in coding pain expressions of SR and OR faces by assessing repetition suppression (RS) of neural activity to pain expressions (Sheng et al., 2016). RS reflects relative attenuation in neural responses to repeated occurrence of a stimulus, and RS of neural activity elicited by two successive stimuli implies engagement of an overlapping neuronal population in the processing of both stimuli (Grill-Spector et al., 2006). Based on ample evidence for decreased neural responses to painful expressions of OR than SR faces (e.g. Xu et al., 2009; Mathur et al., 2010; Sheng and Han, 2012; Sessa et al., 2014; Luo et al., 2015; Zhou and Han, 2021; see Han, 2018 for review), Sheng et al. (2016) recorded electroencephalogram (EEG) to clarify whether coding pain expressions of SR and OR faces engages shared or distinct neuronal activities by examining RS of neural responses to rapidly presented faces. In their design, on each trial, an adaptor face with a painful or neutral expression was followed by a target face with a painful expression. The analysis of event-related potentials (ERPs) to target faces revealed decreased amplitudes of a frontal/central positive activity at 148–208 ms post-stimulus (P2) when a target face was preceded by faces with painful compared to neutral expressions. Most importantly, RS of the P2 amplitude to target faces occurred when adaptor and target faces were of the same race but not when their racial identities were different. These findings suggest that the processing of pain expressions of different races in the P2 time window may recruit distinct neuronal activities.

The current study employed the same paradigm to test the hypotheses that fearful (but not happy) expressions of SR and OR faces are encoded by distinct neural populations at specific stages of neural processes. ERP amplitudes triggered by facial expressions have been used to examine multiple stages of affective processing. For example, the face-sensitive N170 recorded at the lateral occipitotemporal electrodes and the frontal positive activity (P2) respond with enlarged amplitudes to both negative (e.g. fear, anger) and positive (i.e. happy) expressions compared to neutral faces (e.g. Williams et al., 2006; Luo et al., 2010; Calvo et al., 2013; see Hinojosa et al., 2015; Calvo and Nummenmaa, 2016 for review), suggesting early affective processing by differentiating emotional and nonemotional expressions. The amplitude of a following negative wave at 200–350 ms post-stimulus over the frontal/central region (N2) differentiates between positive and negative expressions (e.g. Williams et al., 2006; Luo et al., 2010), reflecting a further refined discrimination of perceived emotional states (Calvo and Nummenmaa, 2016). The amplitudes of two long-latency positive waves after 300 ms

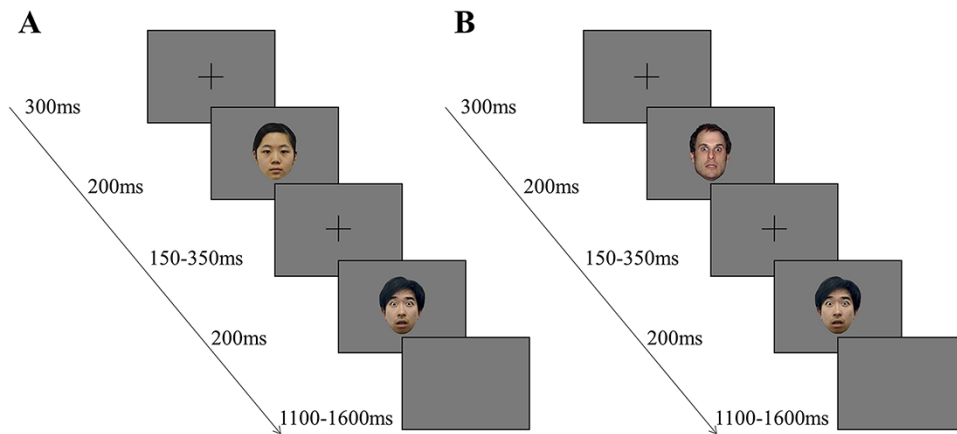


Fig. 1. Illustrations of stimuli and procedure in Experiment 1. A target face with fearful expression was preceded by an adaptor face with either neutral (A) or fearful (B) expressions. The adaptor and target faces were of the same race but of different genders in (A) and were of different races but of the same gender in (B).

post-stimulus are also modulated by facial expressions [e.g. P3 and late positive potential (LPP), Leppänen et al., 2007; Luo et al., 2010], possibly relating to expression categorization (Calvo and Nummenmaa, 2016).

To test whether RS of ERPs to fearful or happy expressions is modulated by perceived race, we presented participants with an adaptor face with an emotional expression (fearful in Experiment 1 and happy in Experiment 2) or a neutral expression on each trial. The adaptor face was followed by a target face with a fearful expression (in Experiment 1) or a happy expression (in Experiment 2). The adaptor and target faces were of the same or different race or gender, as illustrated in Figure 1. Participants were asked to judge whether the adaptor and target faces were of the same gender. This task decreased attention to racial identity of adaptor and target faces and allowed us to test the effect of spontaneous racial categorization of faces on emotion-related RS of brain activities. Our analyses focused on (i) whether ERP amplitudes to fearful/happy target faces decreased following emotional (*vs* neutral) adaptor faces and (ii) whether such RS effects were evident when adaptor and target faces were of the same race but not when they were of difference races. We also analyzed ERPs to adaptor faces to examine the engagement of neural activities in a specific time window in coding race or emotion (i.e. fear and happiness). The current work also included behavioral estimations of explicit and implicit attitudes toward OR and SR faces in order to assess potential individual differences in neural responses to fear/happy expressions related to OR-specific negative attitudes (Hall et al., 2015). Our ERP results suggest that early affective processing of fearful expressions indexed by the P2 amplitude to target faces may engage distinct neural activities for SR and OR faces, whereas late expression categorization (fear/happy *vs* neutral) of SR and OR faces indexed by P3 and LPP may employ shared neural populations.

Methods

Participants

Forty-eight Chinese participants were recruited in Experiment 1, and 54 participants in Experiment 2 were paid volunteers. Five participants in Experiment 1 and 8 participants in Experiment 2 were excluded due to their excessive eye blinks and

head movements during EEG recording. This left 43 participants (mean \pm s.d. = 21.4 ± 2.6 years, 21 males) in Experiment 1 and 46 participants (mean \pm s.d. = 21.6 ± 2.13 years, 21 males) in Experiment 2 for final data analyses. All participants were born and raised in China. All were right-handed, had normal or corrected-to-normal vision and reported no neurological or psychiatric history. Informed consent was obtained from all participants before scanning. This study was approved by the local ethics committee at the School of Psychological and Cognitive Sciences, Peking University.

Stimuli

In Experiment 1, stimuli consisted of digital photographs of faces with neutral or fearful expressions from 16 white models (8 males) and 16 Asian models (8 males). Each model contributed one photograph with a fearful expression and one with a neutral expression. Photographs of white faces were adopted from The NimStim set of facial expressions (Tottenham et al., 2009) and JACFEE (Matsumoto, 1988). Photographs of Chinese faces were taken from volunteers based on explicit criteria of fearful expressions (i.e. eyebrow raised and eyes widened) according to previous research (Ekman and Rosenberg, 1997). The luminance levels of the photographs were matched between white and Asian faces.

To validate the models' facial expressions, we asked an independent sample of participants (46 Chinese, 23 males, mean \pm s.d. = 23.0 ± 2.59 year) to evaluate intensity of fear and attractiveness of each photograph on a 7-point Likert scale (1 = not at all and 7 = extremely fearful or attractive). The participants were also asked to report racial identity of each face and to rate their confidence about racial identity (Asian *vs* white) of each model on a 7-point Likert scale (1 = not at all confident and 7 = extremely confident). The accuracy of racial identity judgment was high ($99\% \pm 0.5\%$) with the mean confidence of 6.5 ± 0.06 . A repeated measure analysis of variance (ANOVA) of rating scores of attractiveness and fear intensity with Race (white *vs* Asian) and Expression (Fear *vs* Neutral) as within-subjects variables showed that, relative to faces with neutral expressions, faces with fearful expressions were rated more fearful but less attractive. However, there was no significant difference in these rating scores between Asian and white faces (see Tables 1 and 2 for statistical details), indicating comparable

Table 1. Results of emotion intensity rating of Asian and white faces in Experiment 1

Fear intensity					
	Asian Neutral	Asian Fearful	White Neutral	White Fearful	
M ± s.d.	1.2 ± 0.51	4.3 ± 0.92	1.2 ± 0.45	4.5 ± 0.91	
Analysis of variance					
	F	df	P	η^2_p	95% CIs
Race	3.018	1,45	0.089	0.063	−0.01, 0.18
Expression	529.404	1,45	<0.001	0.922	2.90, 3.46
Race × Expression	2.431	1,45	0.126	0.051	–

Table 2. Results of attractiveness rating of Asian and white faces in Experiment 1

Attractiveness					
	Asian Neutral	Asian Fearful	White Neutral	White Fearful	
M ± s.d.	1.2 ± 0.51	4.3 ± 0.92	1.2 ± 0.45	4.5 ± 0.91	
Analysis of variance					
	F	df	P	η^2_p	95% CIs
Race	0.065	1,45	0.799	0.001	−0.09, 0.12
Expression	91.010	1,45	<0.001	0.669	−0.94, −0.61
Race × Expression	0.100	1,45	0.753	0.002	–

Table 3. Results of emotion intensity rating of Asian and white faces in Experiment 2

Happy intensity					
	Asian Neutral	Asian Happy	White Neutral	White Happy	
M ± s.d.	1.5 ± 0.60	4.8 ± 0.72	1.5 ± 0.57	4.8 ± 0.75	
Analysis of variance					
	F	df	P	η^2_p	95% CIs
Race	1.965	1,63	0.166	0.030	−0.02, 0.09
Expression	1100.569	1,63	<0.001	0.946	3.12, 3.53
Race × Expression	0.455	1,63	0.502	0.007	–

subjective judgments of facial attractiveness and fear intensity of Asian and white faces.

In Experiment 2, stimuli consisted of digital photographs of faces with neutral or happy expressions from 16 white models (8 males) and 16 Asian models (8 males). Each model contributed one photograph with a happy expression and one with a neutral expression. Photographs of white faces were adopted from The NimStim set of facial expressions (Tottenham *et al.*, 2009) and JACFEE (Matsumoto, 1988). Photographs of Chinese faces were taken from volunteers based on explicit criteria of happy expressions (i.e. eyes wrinkled and mouth drawn back at corners; Ekman and Rosenberg, 1997). The luminance levels of the photographs were matched between white and Chinese faces.

To validate the models' facial expressions, we asked an independent sample of participants (64 Chinese, 32 males, mean ± s.d. = 21.7 ± 2.34 year) to evaluate intensity of happy and attractiveness of each photograph on a 7-point Likert scale (1 = not at all and 7 = extremely happy or attractive). The participants were also asked to report racial identity of each face and to rate their confidence about racial identity (Asian vs white)

of each model on a 7-point Likert scale (1 = not at all confident and 7 = very confident). The accuracy of racial identity judgment was high (98.9 ± 0.2%) with the mean confidence of 6.5 ± 0.06. ANOVA of rating scores of attractiveness and happy intensity with Race (white vs Asian) and Expression (Happy vs. Neutral) as within-subjects variables showed that, relative to faces with neutral expressions, faces with happy expressions were rated happier and more attractive. However, there was no significant difference in these rating scores between Asian and white faces (see Tables 3 and 4 for statistical details), indicating comparable subjective judgments of facial attractiveness and happy intensity of Asian and white faces.

We also estimated possible differences in emotional intensity between the two sets of stimuli in Experiments 1 and 2. We calculated the relative emotional intensity of these two sets of stimuli (i.e. the difference in emotional intensity between fearful/happy and neutral expressions). Independent-samples t-test did not show significant difference in the relative emotional intensity between the two sets of stimuli used in Experiments 1 and 2 { $t(108) = -0.858, P = 0.393, \text{Cohen's } d = -0.164, 95\% \text{ confidence intervals (95\% CIs)} = [-0.473, 0.187]$ }.

Table 4. Results of attractiveness rating of Asian and white faces in Experiment 2

	Attractiveness				
	Asian Neutral	Asian Happy	White Neutral	White Happy	
M ± s.d.	1.5 ± 0.60	4.8 ± 0.72	1.5 ± 0.57	4.8 ± 0.75	
Analysis of variance					
	F	df	P	η^2_p	95% CIs
Race	2.625	1,63	0.110	0.040	−0.02, 0.15
Expression	19.692	1,63	<0.001	0.238	0.19, 0.50
Race × Expression	0.520	1,63	0.474	0.008	–

Procedure

EEG session. In both Experiments 1 and 2, each trial started with fixation cross in the center of a grey background for 500 ms. An adaptor face was then displayed for 200 ms followed by a fixation cross with a duration varying randomly from 150 to 350 ms. The stimulus duration and interstimulus intervals were similar to those used in previous studies in which reliable RS effects on ERPs to faces were observed (Vizioli et al., 2010; Sheng et al., 2016). Next a target face was presented for 200 ms and followed by a blank screen with a duration varying randomly from 1100 to 1600 ms (see Figure 1). Each face subtended a visual angle of $4.0^\circ \times 5.0^\circ$ at a view distance of 80 cm.

In Experiment 1, adaptor faces were selected pseudo-randomly from all 64 faces with fearful or neutral expressions, whereas target faces were selected pseudo-randomly from 32 faces with fearful expressions. In Experiment 2, adaptor faces were selected pseudo-randomly from all 64 faces with happy or neutral expressions, whereas target faces were selected pseudo-randomly from 32 faces with happy expressions. An adaptor and a target on each trial were always different in face identity.

In both Experiments 1 and 2, there were 8 blocks of 128 trials. In each block, half trials consisted of an adaptor and a target with the same expressions (i.e. a fear adaptor followed by a fear target in Experiment 1 and a happy adaptor followed by a happy target in Experiment 2) and half trials consisted of a neutral adaptor followed by a fearful (Experiment 1) or a happy target (Experiment 2). The adaptor and target faces were of the same race for half trials and of different races for other trials. On each trial, participants were asked to judge whether the adaptor and target faces were of the same gender by pressing 1 of 2 keys. The adaptor and target faces were of the same gender for half trials and of different gender for other trials. The relationship between yes/no responses and response keys was counterbalanced across participants.

Behavior session. After EEG recording, participants were asked to report their explicit attitude toward each face by rating a 9-point Likert-type scale (1 = not at all and 9 = extremely) regarding the following question: How much do you like him/her?

We employed a race version of Implicit Association Test (IAT, Greenwald et al., 1998) to assess participants' implicit attitudes toward Asian and White faces. A different set of 10 Asian and 10 white faces (half males) with neutral expressions (Sheng et al., 2016) were used in the IAT. In a block of 20 practicing trials and a block of 40 testing trials, participants were asked to categorize Asian faces/positive words (i.e. joy, love, peace, wonderful, pleasure, glorious, laugh and happy) with one key and white faces/negative words (i.e. agony, terrible, horrible, nasty, evil, awful, failure and hurt) with another key. In another block of 20 practicing trials and a block of 40 testing trials, participants

responded to Asian faces/negative words with one key and white faces/positive words with another key. The IAT was conducted using the software Inquisit 3.0. According to established algorithm (Greenwald et al., 2009), the difference in response speeds between the 2 types of blocks was calculated as an index of racial bias in attitude, namely D score. A D score larger than 0 represents that, compared with outgroup faces, ingroup faces are associated with positive rather than negative attitude, whereas a D score smaller than 0 represents negative rather than positive attitude toward ingroup faces compared with outgroup faces.

EEG acquisition and data analysis

EEG was continuously recorded with Brain Vision Recorder (Brain Products, GmbH) using 64 Ag-AgCl scalp electrodes placed according to the International 10-20 system and referenced online to the FCz electrode. Eye blinks and vertical eye movements were monitored with electrodes located below the right eye. The EEG was amplified (band pass 0.1–1000 Hz) and digitized at a sampling rate of 500 Hz. The EEG was analyzed with Brain Vision Analyzer 2.0 (Brain Products, GmbH). EEG data was re-referenced offline to the average of the left and right mastoid electrodes and then filtered with a band pass from 0.5 to 40 Hz. Eye-movement artifacts were corrected by conducting an independent component analysis. The ERPs in each condition were averaged separately offline with an epoch beginning 200 ms before stimulus onset and continuing for 1200 ms. Trials contaminated by noise exceeding $\pm 70 \mu\text{V}$ at any electrode following adaptor or target faces or response errors were excluded from average. This resulted in 91 ± 14 trials accepted per condition for each participant in Experiment 1 and 90 ± 15 trials accepted per condition for each participant in Experiment 2. The baseline for ERP measurements was the mean voltage of a 200 ms prestimulus interval, and the latency was measured relative to the stimulus onset.

The RS effect related to fear (or happy) expression was defined as decreased ERP amplitudes to target faces, which were preceded by fearful (or happy) vs neutral adaptors. We were particularly interested in the RS effects when the adaptor and target faces were of the same race or different races. ANOVAs were conducted on the amplitudes of N1, P2, N170, P3 and LPP components and behavioral performances [i.e. reaction time (RT) and response accuracy] with Adaptor Race (own-race vs OR), Adaptor Expression (fearful vs neutral in Experiment 1; happy vs neutral in Experiment 2) and Target Race (own-race vs OR) as within-subjects variables. To test our hypotheses by decreasing familywise error rate (Luck and Gaspelin, 2017), we focused on the main effect of Adaptor Expression and the three-way interaction of Adaptor Expression × Adaptor Race × Target Race. The CIs were reported for effect sizes for interaction effects in

ANOVAs (90% CIs of η^2_p) and for the mean difference for other effects (95% CIs of mean difference).

To avoid potential significant but bogus effects on ERP amplitudes due to multiple comparisons (Luck and Gaspelin, 2017), we calculated the mean amplitudes of N1, P2 and N2 elicited by adaptor faces and the mean amplitude of the P2 to target faces over the frontocentral electrode clusters including F1, F2, Fz, FC1, FC2, C1, C2 and Cz. We also calculated mean amplitude values of target N1 and LPP over the central electrode clusters (FC1, FC2, C1, C2, Cz, CP1, CP2 and CPz), the mean amplitude values of target P3 over the parietal electrode clusters (CP1, CP2, CPz, P1, P2 and Pz) and the mean amplitude of the N170 elicited by adaptor and target faces over lateral occipitotemporal electrodes (P7, P8, PO7 and PO8). To avoid any bias due to the selection of time windows for ERP amplitude measurements, we averaged the ERPs across different conditions and then used the timing and scalp distribution of these averaged ERPs to define the time windows for measurements of mean amplitudes of ERP components. For example, to assess the P2 time window, we averaged ERPs to target faces across all the eight conditions (Adaptor Expression \times Adaptor Race \times Target Race, $2 \times 2 \times 2 = 8$). The peak latency of the P2 component in the averaged ERP was set as the center of the P2 time window, and the peak-latency \pm full width at half maximum (30 ms for the P2 component) was set up as the time window for measuring the mean P2 amplitude.

Results

Experiment 1

Behavioral results. Tables 5 and 6 show mean RTs and response accuracies during EEG recording in Experiment 1. ANOVAs of RTs and accuracies for gender judgments (response accuracies were subjected to arcsine-square-root transformation before ANOVAs) did not show any significant effect, providing no evidence for differences in task difficulty in different conditions. Participants reported liking fear faces less than neutral faces, but the rating scores did not differ statistically between SR

and OR faces (see Table 7). One-sample t-test of IAT D scores failed to show significantly difference between the D score and 0 (see Table 7). Therefore, there was no evidence for any difference in explicit or implicit attitudes toward own-race and OR faces. In addition, we examined whether IAT D scores would predict ERP amplitudes related to race/expression processing and RS effects on ERP amplitudes, but failed to find significant correlations. Therefore, these were not reported in the following.

ERPs to adaptor faces. The grand average ERPs to adaptor faces in Experiment 1 were characterized by a negative wave at 90–120 ms (N1), a positive wave at 124–184 ms (P2), a negative deflection at 204–254 ms (N2) over the frontocentral region and a negative wave at the lateral occipitotemporal electrodes at 140–200 ms (N170) (Figure 2A). The time windows of these components were similar to those reported in previous studies (e.g. N1 in Luo et al., 2010; P2 in Holmes et al., 2005; N170 in Bentin et al., 1996 and N2 in Rossion et al., 1999). ANOVAs of the mean N1 amplitudes to adaptor faces did not show any significant effect ($P_s > 0.2$).

ANOVAs of the mean P2 amplitudes to adaptor faces showed a significant main effect of adaptor race ($F(1,42) = 123.315$, $P < 0.001$, $\eta^2_p = 0.746$, 95% CIs = [1.269, 1.833]) and Adaptor Expression ($F(1,42) = 111.515$, $P < 0.001$, $\eta^2_p = 0.726$, 95% CIs = [0.888, 1.307]), suggesting that white adaptor faces elicited larger P2 amplitude compared with Asian adaptor faces and fearful adaptor faces elicited larger P2 amplitude compared with neutral adaptor faces (Figure 2A). These are consistent with previous findings (Williams et al., 2006; Luo et al., 2010; Sheng and Han, 2012; Calvo et al., 2013). There was a significant interaction of Adaptor Race \times Adaptor Expression ($F(1,42) = 5.235$, $P = 0.027$, $\eta^2_p = 0.111$, 90% CIs = [0.007, 0.264]), as the increased P2 amplitude to fearful vs neutral adaptor faces was more salient for Asian faces ($P < 0.001$, mean difference = 1.278, 95% CIs = [1.022, 1.535]) than for white faces ($P < 0.001$, mean difference = 0.917, 95% CIs = [0.647, 1.187]). These results are consistent with previous fMRI research (e.g. Chiao et al., 2008) and indicate that the P2 amplitude is more sensitive to fearful expressions of SR than OR faces.

Table 5. Results of reaction times (ms) during EEG recording in Experiment 1

Reaction time (M \pm s.d.)						
Adaptor Expression	Asian Targets		White Targets			
	Asian Adaptor	White Adaptor	Asian Adaptor	White Adaptor		
Neutral	683 \pm 75.9	699 \pm 78.0	702 \pm 77.7	707 \pm 82.6		
Fearful	684 \pm 78.7	698 \pm 80.3	704 \pm 80.3	704 \pm 79.0		
Analysis of variance						
Main effects and interactions	F	df	P	η^2_p	95% CIs	
Adaptor Expression	0.021	1,42	0.886	0.000	–31, 35	
Adaptor Expression \times Adaptor Race	1.192	1,42	0.281	0.028	–	
Adaptor Expression \times Target Race	0.128	1,42	0.722	0.003	–	
Adaptor Expression \times Adaptor Race \times Target Race	0.409	1,42	0.526	0.010	–	
Simple effect of Adaptor Expression	Mean difference (Neutral – Fearful)	95% confidence interval for difference	F	df	P	η^2_p
Asian–Asian	–9	–66, 49	0.094	1,42	0.761	0.002
Asian–White	–18	–78, 41	0.384	1,42	0.539	0.009
White–Asian	3	–67, 73	0.006	1,42	0.936	0.000
White–White	34	–27, 94	1.274	1,42	0.265	0.029

Table 6. Results of response accuracy (%) during EEG recording in Experiment 1

Adaptor Expression	Response accuracy (%) (M ± s.d.)					
	Asian Targets		White Targets			
	Asian Adaptor	White Adaptor	Asian Adaptor	White Adaptor		
Neutral	92 ± 1.8	95 ± 2.9	96 ± 2.9	94 ± 4.3		
Fearful	91 ± 2.3	95 ± 3.2	96 ± 2.6	95 ± 4.3		
Analysis of variance						
Main effects and interactions	F	df	P	η^2_p	95% CIs	
Adaptor Expression	1.131	1,42	0.294	0.026	−0.018, 0.006	
Adaptor Expression × Adaptor Race	0.307	1,42	0.583	0.007	–	
Adaptor Expression × Target Race	0.062	1,42	0.805	0.001	–	
Adaptor Expression × Adaptor Race × Target Race	2.525	1,42	0.120	0.057	–	
Simple effect of Adaptor Expression	Mean difference (Neutral – Fearful)	95% confidence interval for difference	F	df	P	η^2_p
Asian–Asian	−0.012	−0.036, 0.012	1.047	1,42	0.312	0.024
Asian–White	0.006	−0.016, 0.028	0.296	1,42	0.590	0.007
White–Asian	−0.003	−0.028, 0.022	0.065	1,42	0.800	0.002
White–White	−0.015	−0.033, 0.002	3.220	1,42	0.080	0.071

Note: Response accuracies were subjected to arcsine-square-root transformation before ANOVAs.

Table 7. Results of explicit and implicit attitude estimation in Experiment 1

M ± s.d.	Explicit attitude					
	Asian Neutral	Asian Fearful	White Neutral	White Fearful		
	4.4 ± 0.99	4.1 ± 1.14	4.2 ± 0.99	4.1 ± 1.28		
Analysis of variance						
	F	df	P	η^2_p	95% CIs	
Race	1.258	1,42	0.268	0.029	−0.31, 0.09	
Expression	4.886	1,42	0.033	0.104	−0.42, −0.02	
Race × Expression	3.010	1,42	0.090	0.067	–	
Implicit attitude (IAT D score)						
M ± s.d.	One-sample t	df	P	Cohen's d		
0.078 ± 0.4428	1.145	42	0.259	0.175		

ANOVAs of the N2 amplitudes to adaptor faces showed a significant main effect of Adaptor Race ($F(1,42) = 192.435, P < 0.001, \eta^2_p = 0.821, 95\% \text{ CIs} = [1.691, 2.267]$) and Adaptor Expression ($F(1,42) = 28.799, P < 0.001, \eta^2_p = 0.407, 95\% \text{ CIs} = [0.401, 0.883]$), suggesting that white adaptor faces elicited smaller N2 amplitude compared with Asian adaptor faces and fearful adaptor faces elicited smaller N2 amplitude compared to neutral adaptor faces. However, these effects did not differ significantly between Asian and white faces ($P = 0.360$). ANOVAs of the N170 mean amplitudes revealed significant main effects of Adaptor Race ($F(1,42) = 7.115, P = 0.011, \eta^2_p = 0.145, 95\% \text{ CIs} = [-0.429, -0.059]$) and Adaptor Expression ($F(1,42) = 16.814, P < 0.001, \eta^2_p = 0.286, 95\% \text{ CIs} = [-0.507, -0.173]$), as the N170 was enlarged by fearful compared to neutral expressions and by white than Asian faces (Figure 2B).

ERPs to target faces. The grand average ERPs to target faces in Experiment 1 were characterized by the central N1 (90–130 ms)

and frontocentral P2 (140–200 ms), the occipitotemporal N170 (140–200 ms), the parietal P3 (290–340 ms) and the central LPP (344–524 ms) (Figure 4). ANOVAs of the mean amplitudes of N1 (Adaptor Expression: $P = 0.236$; Adaptor Expression × Adaptor Race × Target Race: $P = 0.514$) did not show any significant RS effects or three-way interactions.

ANOVAs of the mean N170 amplitudes to target faces failed to show significant interaction of Adaptor Expression × Adaptor Race × Target Race at either the right (P8 and PO8: $P = 0.636$) or left (P7 and PO7: $P = 0.082$) temporal electrodes. However, using a linked mastoid reference may limit the possibility to observe ERP components to face stimuli (Joyce and Rossion, 2005) and might mask the effects on these components that our work focused on. Therefore, we reanalyzed the N170 amplitudes by re-referencing EEG to the common average that is created by averaging together signals from all recorded scalp electrodes (Picton et al., 2000). The results revealed a significant main effect of Adaptor Expression ($F(1,42) = 6.519, P = 0.014, \eta^2_p = 0.134, 95\% \text{ CIs} = [-0.389, -0.046]$) at electrodes P8 and PO8, suggesting that

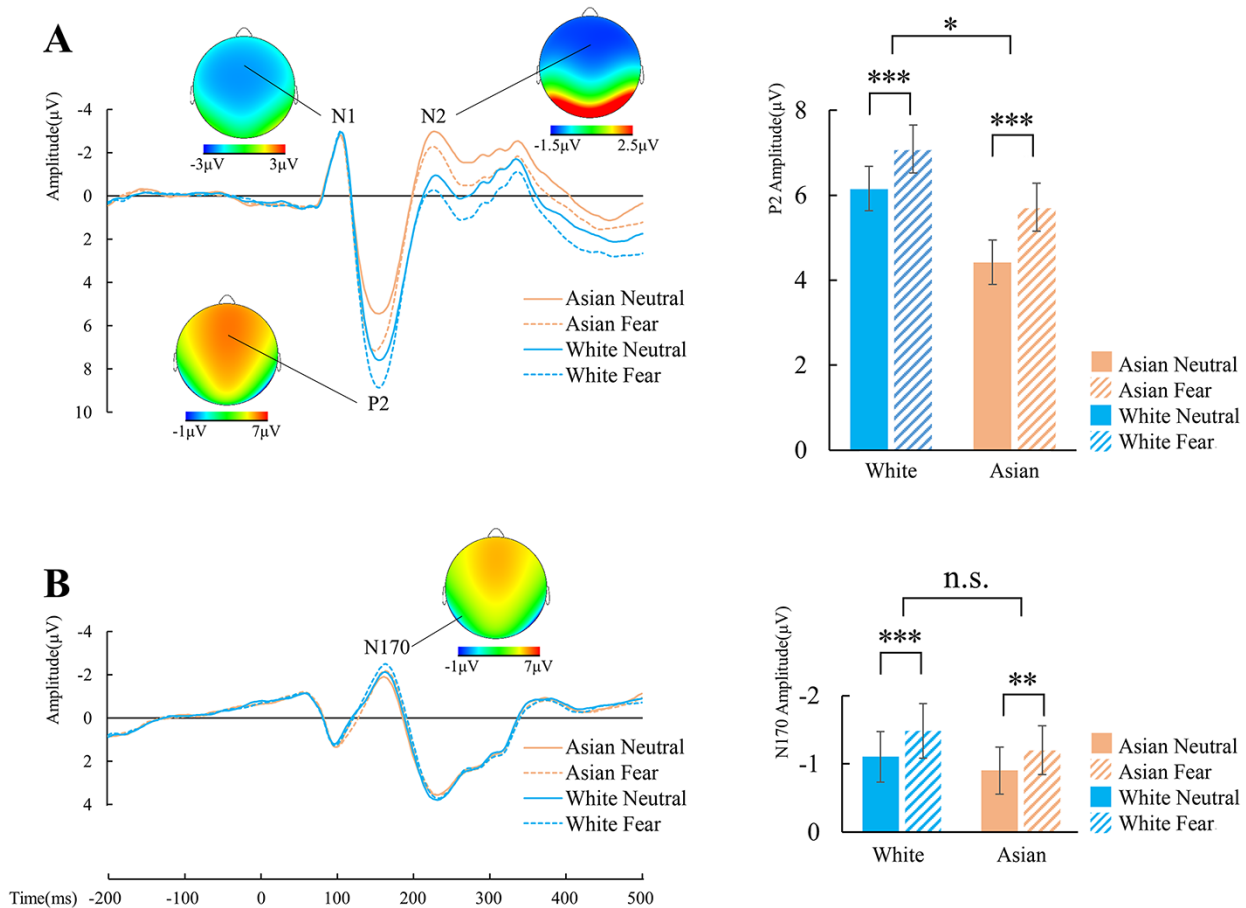


Fig. 2. Illustration of ERPs to adaptor faces in Experiment 1. The left panels show waveforms and voltage topographies of the N1, P2 and N2 (A) and N170 (B). The right panels show the mean amplitudes of the P2 and N170. ** $p < 0.01$; *** $p < 0.001$

the N170 amplitudes to fearful target faces were decreased when preceded by adaptor faces with fearful vs neutral expressions (see Supplementary Figure S1). There was also a significant interaction of Adaptor Expression \times Adaptor Race \times Target Race ($F(1,42) = 9.332$, $P = 0.004$, $\eta^2_p = 0.182$, 90% CIs = [0.037, 0.341]), indicating larger RS effects on the N170 amplitudes to target faces when preceded by SR compared to OR faces. Post hoc pairwise comparisons further revealed that RS of the N170 amplitude to Asian target faces was significant when the target faces were preceded by Asian adaptors ($P = 0.008$, mean difference = -0.419 , 95% CIs = [-0.721 , -0.116]) but not when preceded by white adaptors ($P = 0.228$, mean difference = -0.174 , 95% CIs = [-0.460 , 0.113]). In contrast, the RS effect on the N170 amplitudes to white target faces was significant when the target faces were preceded by white adaptors ($P = 0.004$, mean differences = -0.347 , 95% CIs = [-0.576 , -0.118]) but not when preceded by Asian adaptors ($P = 0.471$, mean differences = 0.070 , 95% CIs = [-0.124 , 0.263]), see Supplementary Figure S1). Analyses of the P2 amplitudes after re-referencing EEG to the common average replicated the results reported in the main text (see Supplementary Figure S2).

ANOVAs of the P2 amplitudes to target faces showed a significant main effect of adaptor expression ($F(1,42) = 6.826$, $P = 0.012$, $\eta^2_p = 0.140$, 95% CIs = [0.061, 0.474]), indicating that the P2 amplitudes to fear target faces were decreased when preceded by adaptor faces with fear vs neutral expressions (Figure 3A). Interestingly, this RS effect on the P2 amplitude was

further quantified by a significant interaction of Adaptor Expression \times Adaptor Race \times Target Race ($F(1,42) = 8.636$, $P = 0.005$, $\eta^2_p = 0.171$, 90% CIs = [0.031, 0.329]). Post hoc pairwise comparisons confirmed that RS of the P2 amplitude to Asian target faces was significant when the target faces were preceded by Asian adaptors ($P = 0.021$, mean difference = 0.469 , 95% CIs = [0.074, 0.864]) but not when preceded by white adaptors ($P = 0.773$, mean difference = 0.052 , 95% CIs = [-0.311 , 0.415]). In contrast, the RS effect on the P2 amplitudes to white target faces was significant when the target faces were preceded by white adaptors ($P = 0.005$, mean differences = 0.485 , 95% CIs = [0.153, 0.818]) but not when preceded by Asian adaptors ($P = 0.657$, mean differences = 0.064 , 95% CIs = [-0.223 , 0.351], Figure 3A). These results indicate that RS of the P2 amplitude to fearful expressions occurred only when adaptor and target faces were of the same race.

ANOVAs of both the mean P3 and LPP amplitudes to target faces showed significant main effects of Adaptor Expression (P3: $F(1,42) = 69.246$, $P < 0.001$, $\eta^2_p = 0.622$, 95% CIs = [0.669, 1.097]; LPP: $F(1,42) = 42.214$, $P < 0.001$, $\eta^2_p = 0.501$, 95% CIs = [0.403, 0.767]), suggesting decreased P3 and LPP amplitudes by fearful vs neutral expressions of adaptor faces (Figure 3B). RS of the P3 amplitude did not show a significant three-way interaction effect ($P = 0.078$). RS of the LPP amplitude was further quantified by a significant interaction of Adaptor Expression \times Adaptor Race \times Target Race ($F(1,42) = 4.351$, $P = 0.043$, $\eta^2_p = 0.094$, 90% CIs = [0.002, 0.243]), suggesting that RS of the LPP amplitude to

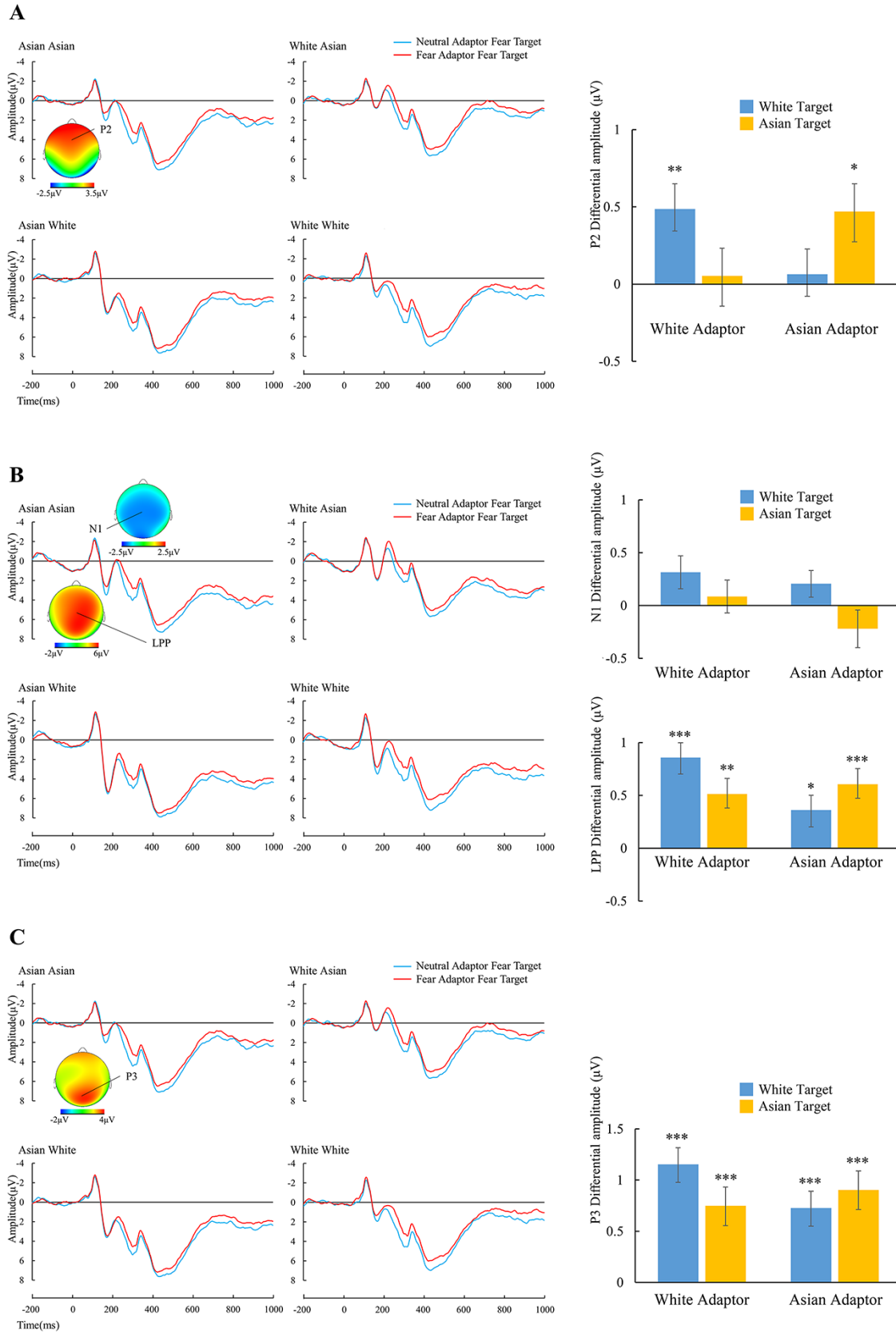


Fig. 3. Illustration of ERPs to target faces in Experiment 1. ERPs to target faces are shown at the frontocentral electrodes (A), the central electrode (B) and the parietal electrodes (C). The left panels show waveforms and voltage topographies of each component, and the right panels show the mean amplitudes of each component. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

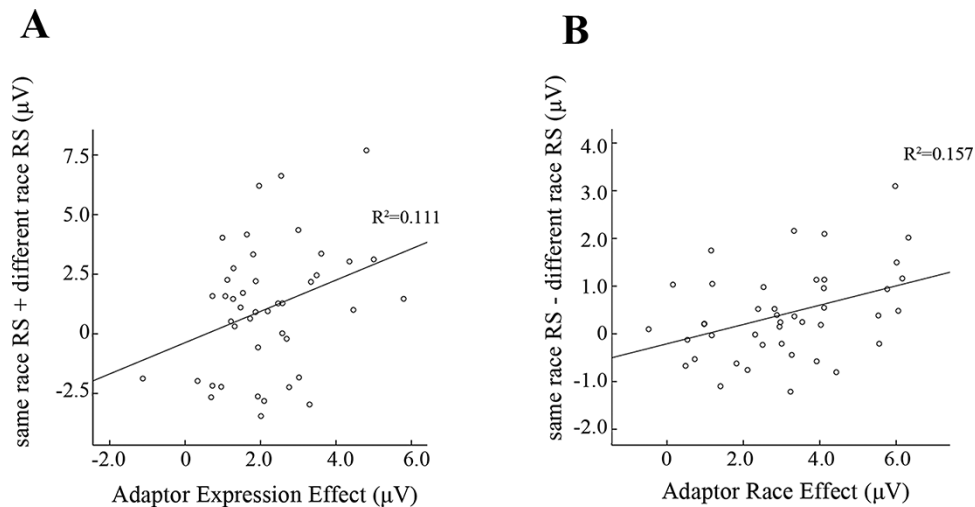


Fig. 4. Results of correlation analyses. (A) A significant correlation between the expression effect on the P2 amplitude to adaptor faces and RS of the P2 amplitude to target faces. (B) A significant correlation between the race effect on the P2 amplitude to adaptor faces and the same-race over other-race advantage in RS of the P2 amplitude to target faces.

fearful expressions was stronger when adaptor and target faces were of the same race than of different races. However, post hoc pairwise comparisons confirmed RS of the LPP amplitude in all conditions (Asian–Asian: $P < 0.001$, mean differences = 0.606, 95% CIs = [0.337, 0.874]; white–white: $P < 0.001$, mean differences = 0.859, 95% CIs = [0.576, 1.143]; Asian–white: $P = 0.028$, mean differences = 0.361, 95% CIs = [−0.042, 0.681]; white–Asian: $P = 0.001$, mean differences = 0.513, 95% CIs = [−0.215, 0.812]). These results suggest that RS of the LPP amplitude to fearful expressions occurred regardless of whether adaptor and target faces were of the same race than of different races.

We conducted correlation analyses to test whether the effect of fearful expression on ERPs to adaptor faces predicted the RS effect on ERPs to target faces. There was a significant correlation between the expression effect on the P2 amplitude to adaptor faces and the RS effect on the P2 amplitude to target faces (collapsing the RS effects when adaptor and target faces were of the same race and of different races) (Pearson's correlation, $r = 0.369$, $P = 0.009$, $R^2 = 0.111$, Figure 4A). This result suggests that individuals with greater P2 response to fearful adaptor faces showed larger RS of the P2 response to fearful target faces. In addition, there was a significant correlation between the race effect on the P2 amplitude to adaptor faces and the race-dependent RS effect on the P2 amplitude to target faces (the RS effect when adaptor and target faces were of the same race minus the RS effect when adaptor and target faces were of different races) (Pearson's correlation, $r = 0.333$, $P = 0.029$, $R^2 = 0.157$, Figure 4B). This finding suggests that individuals with greater neural sensitivity to adaptor races in the P2 time window showed larger SR over OR advantage in RS of neural responses to target faces.

Experiment 2

Behavioral results. Tables 8 and 9 show the mean RTs and response accuracies during EEG recording in Experiment 2. ANOVAs of RTs during gender judgments failed to show any significant effect. ANOVAs of accuracies during gender judgments (response accuracies were subjected to arcsine-square-root transformation before ANOVAs) only showed a significant Adaptor Expression main effect, as participants responded

slightly more accurately for neutral than happy adaptor faces. Participants reported liking happy faces more than neutral faces, but these rating scores did not differ significantly between SR and OR faces (see Table 10). One-sample t-test of IAT D scores revealed that the D score was not significantly different from 0 (see Table 10). These results suggest no evidence for racial ingroup biases in explicit or implicit attitudes. Similarly, we examined whether IAT D scores would predict ERP amplitudes related to race/expression processing and RS effects on ERP amplitudes, but failed to find significant correlations.

ERPs to adaptor faces. Similarly, the grand average ERPs to adaptor faces in Experiment 2 were also characterized by the frontocentral N1 (94–124 ms) and P2 (132–192 ms), the occipitotemporal N170 (140–200 ms) and the frontocentral N2 (210–260 ms) (Figure 5). ANOVAs of the N1 amplitudes to adaptor faces did not show any significant effect ($P > 0.06$).

ANOVAs of the P2 amplitudes to adaptor faces showed a significant main effect of Adaptor Race ($F(1,45) = 159.479$, $P < 0.001$, $\eta^2_p = 0.780$, 95% CIs = [1.415, 1.953]) and Adaptor Expression ($F(1,45) = 26.990$, $P < 0.001$, $\eta^2_p = 0.375$, 95% CIs = [0.231, 0.523]), suggesting that white vs Asian adaptor faces elicited larger P2 amplitudes and happy vs neutral adaptor faces elicited larger P2 amplitudes. There was no significant interaction of Adaptor Race \times Adaptor Expression ($P = 0.305$). ANOVAs of the N2 amplitudes to adaptor faces showed significant main effects of Adaptor Race ($F(1,45) = 186.2$, $P < 0.001$, $\eta^2_p = 0.805$, 95% CIs = [1.338, 1.801]) and Adaptor Expression ($F(1,45) = 35.45$, $P < 0.001$, $\eta^2_p = 0.441$, 95% CIs = [0.394, 0.796]). White vs Asian Adaptor faces elicited smaller N2 amplitudes and happy vs neutral adaptor faces also elicited smaller N2 amplitudes. There was no significant interaction effect of Adaptor Race \times Adaptor Expression ($P = 0.286$). Similarly, ANOVAs of the N170 amplitudes only revealed significant main effects of Adaptor Race ($F(1,45) = 7.421$, $P = 0.009$, $\eta^2_p = 0.142$, 95% CIs = [−0.401, −0.060]) and Adaptor Expression ($F(1,45) = 18.601$, $P < 0.001$, $\eta^2_p = 0.292$, 95% CIs = [−0.356, −0.129]). The N170 was enlarged by happy compared to neutral expressions and by white than Asian faces. However, there was no significant interaction of Adaptor Race \times Adaptor Expression ($P = 0.126$). Together, these results

Table 8. Results of reaction times (ms) during EEG recording in Experiment 2

Adaptor Expression	Reaction time (M ± s.d.)				Analysis of variance	
	Asian Targets		White Targets			
	Asian Adaptor	White Adaptor	Asian Adaptor	White Adaptor		
Neutral	688 ± 110.5	699 ± 109.0	700 ± 113.1	703 ± 107.1		
Happy	687 ± 110.2	694 ± 108.8	700 ± 110.8	701 ± 107.8		
Main effects and interactions		F	df	P	η^2_p	95% CIs
Adaptor Expression		2.245	1,45	0.141	0.048	–8, 57
Adaptor Expression × Adaptor Race		1.228	1,45	0.274	0.027	–
Adaptor Expression × Target Race		0.655	1,45	0.423	0.014	–
Adaptor Expression × Adaptor Race × Target Race		0.129	1,45	0.722	0.003	–
Simple effect of Adaptor Expression	Mean difference (Neutral – Happy)	95% confidence interval for difference	F	df	P	η^2_p
Asian–Asian	15	–35, 64	0.363	1,45	0.550	0.008
Asian–White	3	–52, 59	0.016	1,45	0.901	0.000
White–Asian	56	–4, 115	3.519	1,45	0.067	0.073
White–White	23	–47, 92	0.428	1,45	0.516	0.009

Table 9. Results of response accuracy (%) during EEG recording in Experiment 2

Adaptor Expression	Response accuracy (%) (M ± s.d.)				Analysis of variance	
	Asian Targets		White Targets			
	Asian Adaptor	White Adaptor	Asian Adaptor	White Adaptor		
Neutral	97 ± 2.9	95 ± 3.5	95 ± 3.5	95 ± 4.3		
Happy	97 ± 2.6	95 ± 3.2	95 ± 4.5	94 ± 5.2		
Main effects and interactions		F	df	P	η^2_p	95% CIs
Adaptor Expression		9.920	1,45	0.003	0.181	0.006, 0.026
Adaptor Expression × Adaptor Race		1.226	1,45	0.274	0.027	–
Adaptor Expression × Target Race		0.007	1,45	0.932	0.000	–
Adaptor Expression × Adaptor Race × Target Race		0.213	1,45	0.647	0.005	–
Simple effect of Adaptor Expression	Mean difference (Neutral – Happy)	95% confidence interval for difference	F	df	P	η^2_p
Asian–Asian	0.012	–0.007,0.031	1.671	1,45	0.203	0.036
Asian–White	0.007	–0.017,0.031	0.369	1,45	0.547	0.008
White–Asian	0.019	–0.002,0.041	3.285	1,45	0.077	0.068
White–White	0.026	0.001,0.051	4.421	1,45	0.041	0.089

Note: Response accuracies were subjected to arcsine-square-root transformation before ANOVAs.

provide no evidence for racial ingroup favoritism in neural responses to happy expressions.

ERPs to target faces. The grand average ERPs to target faces in Experiment 2 were characterized by a negative wave at 98–138 ms (N1) over the frontal/central region and a positive wave at 150–210 ms (P2) over the frontocentral region, which were followed by a positive deflection at 294–344 ms (P3) over the parietal region and a long-latency positivity at 396–576 ms (LPP) over the central region (Figure 6).

ANOVAs of the mean N1 amplitudes to target faces only showed a significant main effect of Adaptor Expression ($F(1,45) = 4.899$, $P = 0.032$, $\eta^2_p = 0.098$, 95% CIs = [0.021, 0.444]),

as the N1 amplitudes to target faces were decreased when preceded by happy vs neutral adaptor faces, indicating RS of early neural responses to happy expressions (Figure 6). However, there was no significant three-way interaction ($P = 0.776$), providing no evidence for modulation of the N1 RS effect by the racial relationship between adaptor and target faces. ANOVA of the N170 amplitude also showed a significant RS effect ($F(1,45) = 20.065$, $P < 0.001$, $\eta^2_p = 0.308$, 95% CIs = [0.137, 0.361]), as the N170 amplitudes decreased to target faces when preceded by happy vs neutral adaptor faces. However, similarly, there was no significant three-way interaction ($P = 0.279$). Thus, there was no evidence for modulation of the N170 RS effect by the racial relationship between adaptor and target faces.

Table 10. Results of explicit and implicit attitude estimation in Experiment 2

Explicit attitude					
	Asian Neutral	Asian Happy	White Neutral	Happy	
M ± s.d.	3.5 ± 1.16	5.3 ± 1.21	3.7 ± 1.15	5.3 ± 1.25	
Analysis of variance					
	F	df	P	η^2_p	95% CIs
Race	0.423	1,45	0.519	0.009	-0.10, 0.20
Expression	167.036	1,45	<0.001	0.788	1.44, 1.98
Race × Expression	1.770	1,45	0.190	0.038	-
Implicit attitude (IAT D score)					
M ± s.d.	One-sample t	df	P	Cohen's d	
-0.063 ± 0.5612	-0.758	45	0.453	0.112	

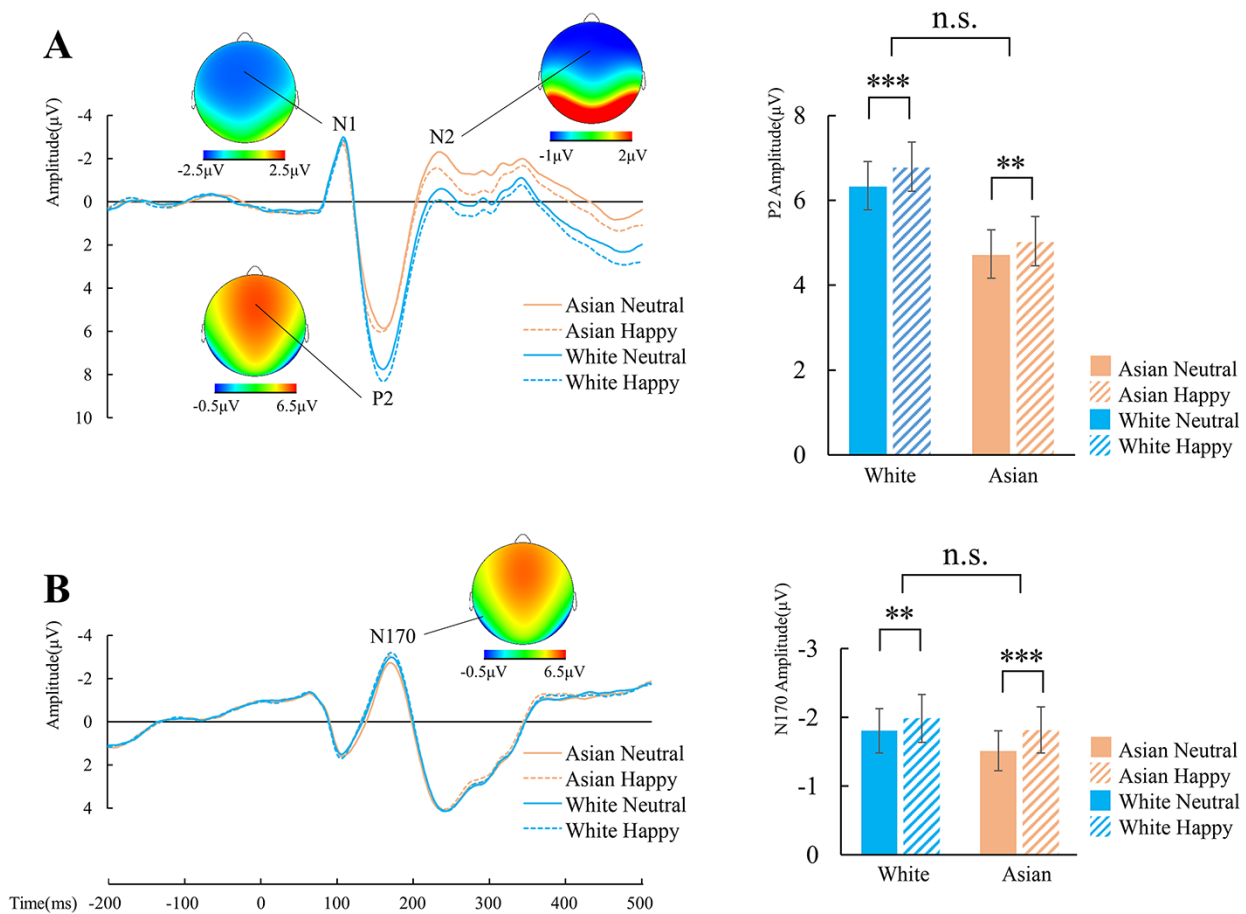


Fig. 5. Illustration of ERPs to adaptor faces in Experiment 2. The left panels show waveforms and voltage topographies of the N1, P2 and N2 (A) and N170 (B). The right panels show the mean amplitudes of the P2 and N170. ** $p < 0.01$; *** $p < 0.001$

ANOVAs of the P2 amplitudes to target faces showed neither significant main effect of Adaptor Expression nor significant interaction of Adaptor Expression × Adaptor Race × Target Race ($P_s > 0.7$). ANOVAs of the P3 and LPP amplitudes to target faces only showed significant main effects of Adaptor Expression (P3: $F(1,45) = 13.490$, $P = 0.001$, $\eta^2_p = 0.231$, 95% CIs = [0.154, 0.529]; LPP: $F(1,45) = 15.486$, $P < 0.001$, $\eta^2_p = 0.256$, 95% CIs = [0.183, 0.566]), suggesting that the P3 and LPP amplitudes to happy

target faces were decreased when preceded by happy vs neutral adaptor faces. However, there was no other significant effect ($P_s > 0.29$), providing no evidence for modulations of P3/LPP RS effects by the racial relationship between adaptor and target faces.

Finally, we assessed how likely to observe a significant three-way interaction effect, given the sample sizes in Experiment 2 and the effect size observed in Experiment 1 using G*Power

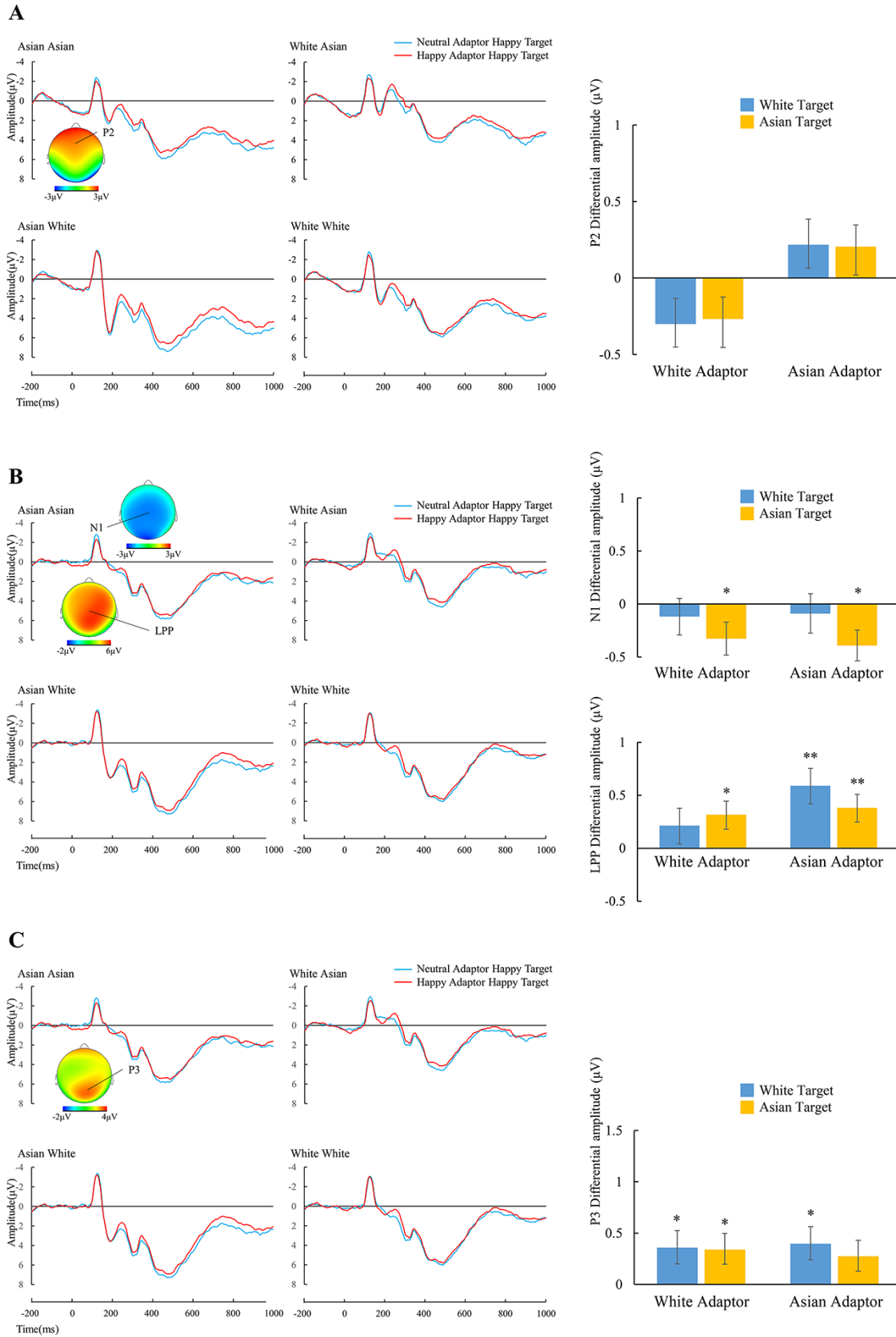


Fig. 6. Illustration of ERPs to target faces in Experiment 2. ERPs to target faces are shown at the frontocentral electrodes (A), the central electrode (B) and the parietal electrodes (C). The left panels show waveforms and voltage topographies of each component, and the right panels show the mean amplitudes of each component. * $p < 0.05$; ** $p < 0.01$

3.1 (Faul et al., 2009). The results showed that given the effect size of the three-way interaction effect on the P2 amplitudes observed in Experiments 1 (i.e. $\eta^2_p = 0.171$), the sample size in Experiment 2 had a power of 0.99 to detect a reliable interaction effect of the P2 amplitude, which is well above conventional recommendations. The results suggest that the absence of the three-way interaction effect in Experiment 2 was not simply due to an underpowered sample size. Instead, the results suggest that the RS effect of happy expression in the P2 time window was not affected by the racial relationship between adaptor and target faces.

Similarly, we reanalyzed the N170 and P2 amplitudes in Experiment 2 by re-referencing EEG to the common average that is created by averaging together signals from all recorded scalp electrodes. The results replicated those reported in the main text (see Supplementary Figures S3 and S4).

Discussion

In two experiments, we tested the hypotheses that the processing of fearful but not happy expressions of SR and OR faces engages distinct neural activities by recording EEG to adaptor and target faces. Our data analyses focused on the RS effects on neural responses to faces with fearful (in Experiment 1) and happy (in Experiment 2) expressions by comparing ERP amplitudes to target faces that were preceded by adaptor faces with fearful/happy or neutral expressions. We were particularly interested in whether RS effects were independent of perceived race of adaptor and target faces. Because the task instruction drew participants' attention to genders of adaptor and target faces, our EEG results revealed how automatic processing of facial expressions is influenced by spontaneous racial categorization of faces. Participants' response speeds and accuracies were not affected by the racial relationship between adaptor and target faces. In addition, both emotional intensity of fearful and happy faces and implicit/explicit attitudes toward SR and OR faces were controlled. Together, these behavioral results suggest that our EEG findings cannot be accounted for by any difference in task difficulty, attitude or intensity of facial expressions between SR and OR faces.

Our ERP results first showed that the P2 and N2 amplitudes to adaptor faces were sensitive to both perceived race and expressions. The P2 amplitude was enlarged by OR than SR faces and by fearful/happy than neutral faces, whereas the N2 amplitude illustrated a reverse pattern of modulations by perceived race and emotion. These results are consistent with previous EEG findings (Williams et al., 2006; Kubota and Ito, 2007; Luo et al., 2010; Sheng and Han, 2012; Calvo et al., 2013; Zhou et al., 2020) and suggest similar time courses of early neural activities in response to perceived race and expressions in the current and previous studies. Moreover, the P2 amplitude in response to adaptor faces showed evidence for interactions between perceived race and expression in Experiment 1 but not in Experiment 2. Similar interaction between perceived race and painful expression was observed in the previous ERP studies of empathy for pain (Sheng and Han, 2012; Sheng et al., 2013; Luo et al., 2018). While these results unraveled early interactions between fearful expression and racial identity on the neural processes of faces, possibly due to early spontaneous categorization of OR faces by race (Han, 2018; Zhou et al., 2020), it remains unclear whether distinct or common neural activities were involved in the processing of fearful and happy expressions of SR and OR faces.

We addressed this issue by examining RS effects on ERP amplitudes to target faces. As expected, two critical findings emerged from ERPs to target faces that support the hypothesis of distinct neuronal activities for the processing of fearful but not happy expressions of SR and OR faces. First, our analyses of ERP amplitudes to target faces uncovered significant RS effects on neural responses to target faces in several time windows. For example, the N1 amplitudes to happy target faces at 98–138 ms were decreased following happy compared to neutral adaptor faces in Experiment 2. The P2 amplitudes to fearful target faces at 140–200 ms were also reduced when preceded by fearful compared to neutral adaptor faces in Experiment 1. In both Experiments 1 and 2, relative to neutral adaptor faces, fearful or happy adaptor faces significantly decreased the P3 and LPP amplitudes to fearful or happy target faces. These results indicate that repetition of two faces with the same (either fearful or happy) expressions inhibit multiple neural processes of the second face in a broad time window in which both early processes of affective states (e.g. emotional vs neutral) and late expression categorization (Calvo and Nummenmaa, 2016) take place. These results highlight that the adaptor and target faces with same expressions (either fearful or happy) shared neural activities for both early processes of affective states (e.g. emotional vs neutral) and late expression categorization.

Second, and most importantly, we found evidence that RS of the P2 amplitude to fearful faces was significantly modulated by the racial relationship between adaptor and target faces. Specifically, RS of the P2 amplitude to fearful target faces was evident when adaptor and target faces were of the same race but not when they were of different races. In consistent with the RS effect on the P2 amplitudes to fearful target faces, the results of our re-reference analyses revealed similar patterns of modulations of the N170 amplitudes to target faces by the racial relationship between adaptor and target faces. These results suggest a specific consequence of a shared neural activity underlying the processing of fearful expressions of faces with the same racial identity and possible distinct neural activities engaged in the processing of fearful expressions of faces with different racial identities. It should be noted that the finding of SR over OR advantage in the RS effect on the P2 amplitude to target faces was observed for both Asian and white faces. This suggests that the SR over OR advantage in the RS effect on neural responses to target faces in the P2 time window could not be produced by some features specific to Asian faces because white and Asian faces showed similar SR over OR advantage of the RS effects.

In Experiment 2, we found a reliable RS effect on the N1 amplitude to happy target faces regardless of whether the adaptor and target faces were of the same race. This is consistent with previous findings of enlarged N1 amplitude to happy than non-happy faces (Calvo et al., 2014) and suggests early neural coding of happy expressions. However, our results did not show effects of consistency of perceived racial identity of adaptor and target faces on the N1 modulation by happy (vs neutral) expressions. Similarly, RS effects on the P3 and LPP amplitudes to neither fearful nor happy target faces were influenced by perceived racial relationships between adaptor and target faces. While our ERP results provide evidence for RS to neural responses to fearful and happy expressions in a wide time window that covers early affective processing (e.g. P2, Williams et al., 2006; Luo et al., 2010; Calvo et al., 2013; Calvo and Nummenmaa, 2016) and late expression categorization (e.g. P3 and LPP, Leppänen et al., 2007; Luo et al., 2010; Calvo and Nummenmaa, 2016), our ERP results

suggest that the consistency of racial identity of adaptor and target faces specifically affected RS of the P2 amplitude to fearful target faces.

Together, the RS effect on the P2 amplitudes to target faces observed in our study supports the hypothesis that distinct neural activities may be engaged during the processing of fearful but not happy expressions of SR and OR faces. Our findings have important implications for understanding the neural processing of facial expressions. The effect of perceived race on RS of the P2 amplitude to fearful faces in the current work is congruent with the previous finding that RS of the P2 amplitude to painful faces was similarly modulated by perceived race of adaptor and target faces (Sheng et al., 2016). Neural responses to perceived painful stimuli to ingroup members or SR individuals are localized to the anterior cingulate and insula and predict motives to help ingroup or SR individuals who suffer, whereas perceived painful stimuli to outgroup members or OR individuals activate the reward system and predict motives of not helping those who suffer (e.g. Hein et al., 2010; Luo et al., 2015). These neuroimaging findings imply that perceived pain in ingroup and outgroup or in SR and OR individuals delivers different social information that in turn trigger different social actions. It is thus not surprising that distinct neuronal populations may be engaged to code painful expressions of SR and OR faces as qualitatively discrepant social signals. Similarly, if fearful expressions of ingroup and outgroup members also deliver different social information (Weisbuch and Ambady, 2008), neural coding of fearful expressions of SR and OR faces with distinct neural activities would facilitate quick and different actions toward these faces. These results highlight the potential effects of social signals delivered by facial expressions on the development of neural strategies for the processing of expressions of SR and OR faces and for taking appropriate adaptive responses quickly to facilitate observers' survival.

Our analyses of the N170 in response to adaptor faces verified the previous finding that faces with either negative (e.g. fearful) or positive (e.g. happy) expressions evoked larger N170 (i.e. more negative) compared to neutral faces (Hinojosa et al., 2015). Although the N170 to happy target faces showed a significant RS effect, this effect did not vary as a function of the consistency of racial identity of adaptor and target faces. However, the N170 in response to fearful target faces exhibited a reliable RS effect only when the adaptor and target faces were of the same race. The N170 is supposed to originate from the superior temporal sulcus or the fusiform gyrus (Sadeh et al., 2010). Our results suggest that the N170 amplitudes to fearful target faces were sensitive to both expressions and perceived racial relationship between adaptor and target faces. These results suggest that the P2 and N170 in response to fearful/happy faces are different regarding sensitivity to racial identities of faces.

The RS effect on the amplitudes of late ERP components (e.g. P3 and LPP) of both fearful and happy faces was not influenced by perceived relationship between the adaptor and target faces either. These results suggest that SR and OR expressions (either fearful or happy) shared neural activities in the late expression categorization (Calvo and Nummenmaa, 2016), which may provide a neural basis of successful cross-cultural recognition of basic emotion expressions.

Previous EEG/MEG research on long-lag incidental repetition of neural responses to fearful, happy and neutral faces observed RS effects as early as 40 ms post-stimulus and extended to 320 ms post-stimulus (Morel et al., 2009). Similarly, the current work showed evidence for RS of neural responses to fearful and happy

faces in a broad time window that covers early affective processing and late expression categorization of faces. These findings leave open new questions. For example, while the effect of racial identity of adaptor and target faces on RS of neural responses to fearful faces was observed only in the P2 time window, similar to that observed for painful expressions (Sheng et al., 2016), the current work was unable to localize the RS effect in the P2 time window due to the limitation of spatial resolution of EEG measures. Future research should localize the brain regions in which distinct neuronal populations are recruited to code fearful/painful expressions of SR and OR faces. An additional issue arising from the current work is how distinct neuronal populations underlying the processing of fearful expressions influence behavioral responses to SR and OR individuals during social interactions. To clarify this question would advance our understanding of the functional role of neural coding of facial expressions in social behavior. Finally, because the current work tested only Chinese participants, future research should employ a cross-race design to compare the RS effects on neural responses to fearful faces in different cultural groups (e.g. Asian and white participants). This is necessary to make a general conclusion that SR and OR faces recruit distinct neuronal activities to encode fearful expressions in different cultural groups.

In conclusion, our ERP findings provide initial evidence that coding fearful expressions of SR and OR faces recruits distinct neuronal populations in an early time window of face processing. By contrast, there was no evidence for distinct neuronal populations underlying the processing of happy expressions of SR and OR faces. Our results support the view of function of facial expressions as important social signals related to survival (Shariff and Tracy, 2011) and advance our understanding of the neural mechanisms underlying the processing of facial expressions from an adaptive perspective. Research along this line helps to comprehend functional roles of other facial expressions in modulating intercultural communications and social interactions in future.

Acknowledgement

We thank the National Center for Protein Sciences at Peking University for assistance with both experiments.

Funding

This work was supported by the National Natural Science Foundation of China (projects 31871134 and 31421003) and Ministry of Science and Technology of China (2019YFA0707103).

Conflict of interest

None declared.

Supplementary data

Supplementary data are available at SCAN online.

References

- Bentin, S., Allison, T., Puce, A., Perez, E., McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, 8(6), 551–65.
- Bourgeois, P., Hess, U. (2008). The impact of social context on mimicry. *Biological Psychology*, 77(3), 343–52.

- Calvo, M.G., Marrero, H., Beltrán, D. (2013). When does the brain distinguish between genuine and ambiguous smiles? An ERP study. *Brain and Cognition*, *81*(2), 237–46.
- Calvo, M.G., Beltrán, D., Fernández-Martín, A. (2014). Processing of facial expressions in peripheral vision: neurophysiological evidence. *Biological Psychology*, *100*, 60–70.
- Calvo, M.G., Nummenmaa, L. (2016). Perceptual and affective mechanisms in facial expression recognition: an integrative review. *Cognition & Emotion*, *30*(6), 1081–106.
- Chiao, J.Y., Iidaka, T., Gordon, H.L., et al. (2008). Cultural specificity in amygdala response to fear faces. *Journal of Cognitive Neuroscience*, *20*(12), 2167–74.
- Cosmides, L., Tooby, J., Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Sciences*, *7*(4), 173–9.
- Darwin, C. (1872). *The Expression of the Emotions in Man and Animal*. London: J. Murray.
- Ekman, P., Friesen, W.V. (1978). *Facial Action Coding System: Investigator's Guide*. Palo Alto, CA: Consulting Psychologists Press.
- Faul, F., Erdfelder, E., Buchner, A., Lang, A. (2009). Statistical power analyses using GPower 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–60.
- Greenwald, A.G., McGhee, D.E., Schwartz, J. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–80.
- Greenwald, A.G., Poehlman, T.A., Uhlmann, E.L., Banaji, M.R. (2009). Understanding and using the implicit association test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*(1), 17–41.
- Grill-Spector, K., Henson, R., Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14–23.
- Hall, W.J., Chapman, M.V., Lee, K.M., et al. (2015). Implicit racial/ethnic bias among health care professionals and its influence on health care outcomes: a systematic review. *American Journal of Public Health*, *105*(12), e60–76.
- Han, S. (2018). Neurocognitive basis of racial ingroup bias in empathy. *Trends in Cognitive Sciences*, *22*(5), 400–21.
- Hein, G., Silani, G., Preuschoff, K., Batson, C.D., Singer, T. (2010). Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron*, *68*(1), 149–60.
- Hinojosa, J.A., Mercado, F., Carretié, L. (2015). N170 sensitivity to facial expression: a meta-analysis. *Neuroscience and Biobehavioral Reviews*, *55*, 498–509.
- Holmes, A., Winston, J.S., Eimer, M. (2005). The role of spatial frequency information for ERP components sensitive to faces and emotional facial expression. *Cognitive Brain Research*, *25*(2), 508–20.
- Iidaka, T., Nogawa, J., Kansaku, K., Sadato, N. (2008). Neural correlates involved in processing happy affect on same race faces. *Journal of Psychophysiology*, *22*(2), 91–9.
- Joyce, C., Rossion, B. (2005). The face-sensitive N170 and VPP components manifest the same brain processes: the effect of reference electrode site. *Clinical Neurophysiology*, *116*(11), 2613–31.
- Kret, M.E., Stekelenburg, J.J., Roelofs, K., Gelder, B.D. (2013). Perception of face and body expressions using electromyography, pupillometry and gaze measures. *Frontiers in Psychology*, *4*, 28.
- Kubota, J.T., Ito, T.A. (2007). Multiple cues in social perception: the time course of processing race and facial expression. *Journal of Experimental Social Psychology*, *43*(5), 738–52.
- Kurzban, R., Tooby, J., Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(26), 15387–92.
- Leppänen, J.M., Kauppinen, P., Peltola, M.J., Hietanen, J.K. (2007). Differential electrocortical responses to increasing intensities of fearful and happy emotional expressions. *Brain Research*, *1166*(1), 103–9.
- Luck, S.J., Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't): how to get significant effects. *Psychophysiology*, *54*(1), 146–57.
- Luo, S., Li, B., Ma, Y., Zhang, W., Rao, Y., Han, S. (2015). Oxytocin receptor gene and racial ingroup bias in empathy-related brain activity. *Neuroimage*, *110*, 22–31.
- Luo, S., Han, X., Du, N., Han, S. (2018). Physical coldness enhances racial in-group bias in empathy: electrophysiological evidence. *Neuropsychologia*, *116*(Pt A), 117–25.
- Luo, W., Feng, W., He, W., Wang, N., Luo, Y. (2010). Three stages of facial expression processing: ERP study with rapid serial visual presentation. *NeuroImage*, *49*(2), 1857–67.
- Mathur, V.A., Harada, T., Lipke, T., Chiao, J.Y. (2010). Neural basis of extraordinary empathy and altruistic motivation. *Neuroimage*, *51*(4), 1468–75.
- Matsumoto, D., Ekman, P. (1988). *Japanese and Caucasian facial expressions of emotion (JACFEE)*. [Slides]. San Francisco: Intercultural and Emotion Research Laboratory, Department of Psychology, San Francisco State University.
- Morel, S., Ponz, A., Mercier, M., Vuilleumier, P., George, N. (2009). EEG-MEG evidence for early differential repetition effects for fearful, happy and neutral faces. *Brain Research*, *1254*, 84–98.
- Nelson, N.L., Russell, J.A. (2013). Universality revisited. *Emotion Review*, *5*(1), 8–15.
- Öhman, A., Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological Review*, *108*(3), 483–522.
- Picton, T.E., Bentin, S., Berg, P., et al. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology*, *37*(2), 127–52.
- Preuschoft, S., van Hooff, J.A.R.A.M. (1997). The social function of “smile” and “laughter”: variations across primate species and societies. In: Segerstål, U., Molnar, P., editors. *Nonverbal Communication: Where Nature Meets Culture*, Mahwah, NJ: Erlbaum, 171–90.
- Rossion, B., Campanella, S., Gomez, C.M., et al. (1999). Task modulation of brain activity related to familiar and unfamiliar face processing: an ERP study. *Clinical Neurophysiology*, *110*(3), 449–62.
- Sadeh, B., Podlipsky, I., Zhdanov, A., Yovel, G. (2010). Event-related potential and functional MRI measures of face-selectivity are highly correlated: a simultaneous ERP-fMRI investigation. *Human Brain Mapping*, *31*(10), 1490–501.
- Sessa, P., Meconi, F., Castelli, L., Dell'Acqua, R. (2014). Taking one's time in feeling other-race pain: an event-related potential investigation on the time-course of cross-racial empathy. *Social Cognitive and Affective Neuroscience*, *9*, 454–63.
- Shariff, A.F., Tracy, J.L. (2011). What are emotion expressions for? *Current Directions in Psychological Science*, *20*(6), 395–9.
- Sheng, F., Liu, Y., Zhou, B., Zhou, W., Han, S. (2013). Oxytocin modulates the racial bias in neural responses to others' suffering. *Biological Psychology*, *92*(2), 380–6.
- Sheng, F., Han, X., Han, S. (2016). Dissociated neural representations of pain expressions of different races. *Cerebral Cortex*, *26*(3), 1221–33.

- Sheng, F., Han, S. (2012). Manipulations of cognitive strategies and intergroup relationships reduce the racial bias in empathic neural responses. *Neuroimage*, *61*(4), 786–97.
- Sorce, J., Emde, R., Campos, J., Klinnert, M. (1985). Maternal emotional signaling: its effect on the visual cliff behavior of 1-year-olds. *Developmental Psychology*, *21*(1), 195–200.
- Tottenham, N., Tanaka, J.W., Leon, A.C., et al. (2009). The Nim-Stim set of facial expressions: judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–9.
- van der Schalk, J., Fischer, A., Doosje, B., et al. (2011). Convergent and divergent responses to emotional displays of ingroup and outgroup. *Emotion*, *11*(2), 286–98.
- 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), 20081–6.
- Vytal, K., Hamann, S. (2010). Neuroimaging support for discrete neural correlates of basic emotions: a voxel-based meta-analysis. *Journal of Cognitive Neuroscience*, *22*(12), 2864–85.
- Weisbuch, M., Ambady, N. (2008). Affective divergence: automatic responses to others' emotions depend on group membership. *Journal of Personality and Social Psychology*, *95*(5), 1063–79.
- Williams, L.M., Palmer, D., Liddell, B.J., Song, L., Gordon, E. (2006). The 'when' and 'where' of perceiving signals of threat versus non-threat. *Neuroimage*, *31*(1), 458–67.
- Xu, X., Zuo, X., Wang, X., Han, S. (2009). Do you feel my pain? Racial group membership modulates empathic neural responses. *Journal of Neuroscience*, *29*(26), 8525–9.
- Zhou, Y., Gao, T., Zhang, T., et al. (2020). Neural dynamics of racial categorization predicts racial bias in face recognition and altruism. *Nature Human Behaviour*, *4*(1), 69–69.
- Zhou, Y., Han, S. (2021). Neural dynamics of pain expression processing: alpha-band synchronization to same-race pain but desynchronization to other-race pain. *Neuroimage*, *224*, 117400.