(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

Biological Psychology 92 (2013) 380–386

Contents lists available at SciVerse ScienceDirect

Biological Psychology

journal homepage: www.elsevier.com/locate/biopsycho

Oxytocin modulates the racial bias in neural responses to others' suffering

Feng Sheng[a,b], Yi Liu[a], Bin Zhou[c], Wen Zhou[c], Shihui Han[a,*]

[a] Department of Psychology, Peking University, Beijing, PR China
[b] Guanghua School of Management, Peking University, Beijing, PR China
[c] Institute of Psychology, Chinese Academy of Sciences, Beijing, PR China

ARTICLE INFO

Article history:
Received 2 July 2012
Accepted 29 November 2012
Available online xxx

Keywords:
Empathy
Oxytocin
ERP
Racial bias

ABSTRACT

The intergroup relationship between a perceiver and a target person influences empathic neural responses to others' suffering, which are increased for racial in-group members compared to out-group members. The current study investigated whether oxytocin (OT), a neuropeptide that has been linked to empathic concern and in-group favoritism, contributes to the racial bias in empathic neural responses. Event-related brain potentials were recorded in Chinese male adults during race judgments on Asian and Caucasian faces expressing pain or showing a neutral expression after intranasal self-administration of OT or placebo. A fronto-central positive activity at 128–188 ms (P2) was of larger amplitude in response to the pain expressions compared with the neutral expressions of racial in-group members but not of racial out-group members. OT treatment increased this racial in-group bias in neural responses and resulted in its correlation with a positive implicit attitude toward racial in-group members. Our findings suggest that OT interacts with the intergroup relationship to modulate empathic neural responses to others' suffering.

1. Introduction

Empathy is the ability to understand and share the emotional states of others, and it plays a key role in prosocial behavior (Batson, 1998; de Waal, 2008). Recent neuroimaging research has resulted in increased interest in the neural mechanisms underlying empathy. One line of research has shown that viewing others in pain activates the midcingulate cortex, the anterior insula, and the sensorimotor cortex (Singe et al., 2004; Avenanti et al., 2005; Gu and Han, 2007; Saarela et al., 2007; Han et al., 2009; Ma et al., 2011). Because these brain regions are engaged in the first-hand experience of pain (Rainville et al., 1997; Wager et al., 2004), it has been proposed that empathy for pain shares the neural mechanisms with the first-hand experience of pain.

Recent studies have shown that empathic neural responses to perceived pain in others are strongly shaped by the social relationship between an observer and a target person. Using functional magnetic resonance imaging (fMRI), Xu et al. (2009) scanned Chinese and Caucasian participants while they watched video segments showing a Chinese or a Caucasian model receiving painful (needle penetration) versus non-painful (Q-tip touch) stimulation. In both Chinese and Caucasian participants, the empathic neural activity in the midcingulate region was much stronger when viewing painful stimulation applied to same-race models compared to other-race models. A subsequent event-related brain potential (ERP) study found that a positive activity at 128–188 ms over the frontal/central brain regions (P2) increased in response to pain expressions versus neutral expressions and that the P2 empathic responses were significantly reduced toward racial out-group faces compared to racial in-group faces (Sheng and Han, 2012). Source estimation suggested that the P2 component may arise from the midcingulate region, which is consistent with the previous fMRI findings (Xu et al., 2009). Avenanti et al. (2010) found that observing the pain of racial in-group but not racial out-group models inhibited the onlookers' sensorimotor activity, as if they were receiving painful stimulation. These results indicate that empathic neural responses to others' suffering are modulated by the intergroup relationships between a perceiver and a target person and that neural responses to others' pain are stronger for racial in-group members compared with racial out-group members.

While the prior studies have demonstrated a racial bias in empathic neural responses, it remains unclear how neurobiological factors may contribute to this effect. Oxytocin (OT) is a neuropeptide that is important for the maintenance of social groups and the development of trust among in-group members (see De Dreu, 2012 for a review). It has been demonstrated that intranasally administered OT, versus placebo, can enhance the behavioral index of emotional empathy in response to positive and negative stimuli (Hurlemann et al., 2010). OT can improve performance during inference of others' emotion (Domes et al., 2007), suggesting that OT

* Corresponding author at: Department of Psychology, Peking University, 5 Yiheyuan Road, Beijing 100871, PR China. Tel.: +86 10 6275 9138; fax: +86 10 6276 1081.
E-mail address: shan@pku.edu.cn (S. Han).

0301-0511/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.biopsycho.2012.11.018

may up-regulate empathic concern for others. However, the effects of OT on social cognition and prosocial behavior are influenced by the social context (Bartz et al., 2011). OT promoted trust or cooperation with in-group members but not with out-group members (De Dreu et al., 2010, 2011). Thus it is likely that OT may improve empathic neural responses specifically to racial in-group members rather than function as a general facilitator of empathy.

The current study tested this hypothesis using a randomized double-blind within-subjects placebo-controlled design. From male adult subjects, we recorded ERPs to racial in-group and out-group faces expressing pain or showing a neutral expression. The facial stimuli were adopted from our previous research (Sheng and Han, 2012). Intranasally administered OT or a placebo was administered to participants on two separate days before the ERPs were recorded. Our previous work showed that the racial bias in fronto-central P2 empathic neural responses was modulated by manipulation of cognitive strategies and intergroup relationships. Specifically, the racial bias was reduced by enhanced attention to individuals' emotions and by including other-race individuals in one's own team for competitions (Sheng and Han, 2012). The current study investigated whether and how P2 empathic neural responses are modulated by OT treatment. If OT plays a role in racial bias in empathy, the in-group bias in the P2 effect observed in Sheng and Han (2012) should be increased by OT, as compared to the placebo treatment. Because Avenanti et al. (2010) found that a deficit in empathic reactivity to racial out-group members was greater in the onlookers who exhibited stronger implicit racial bias in their attitudes, we measured participants' implicit attitudes toward racial in-group and out-group faces. We used the Implicit Association Test (Greenwald et al., 1998) to assess whether OT affects the association between racial bias in empathic neural responses and participants' implicit racial attitudes. We also measured participants' empathy traits using the Interpersonal Reactivity Index (IRI, Davis, 1983) in order to estimate whether individuals with stronger empathy traits would show greater empathic neural responses and whether the effect of OT on empathic neural responses varies as a function of an individual's empathy traits.

2. Method

2.1. Participants

Sixteen Chinese male adults aged 18–26 years ($M = 21.88$, SD = 2.06) participated in this study as paid volunteers. The exclusion criteria included self-reported medical or psychiatric illness and use of medication. Informed consent was obtained prior to participation. The details of the experimental procedure were explained to the participants before the study. This study was approved by a local ethics committee.

2.2. Stimuli and procedure

The stimuli were 32 faces from 16 Asian models and 32 faces from 16 Caucasian models used in our previous work (Sheng and Han, 2012), as illustrated in Fig. 1. One-half of the faces expressed pain, and one-half showed a neutral expression. Emotional intensity, facial attractiveness, and luminance levels were matched between the Asian and Caucasian faces (Sheng and Han, 2012).

In a double-blind, placebo-controlled, within-subjects design, participants took part in two electroencephalograph (EEG) sessions, between 7 and 16 days apart ($M = 10.4$, SD = 3.07). Before each EEG session, 32 IU OT or placebo (containing the active ingredients except for the neuropeptide) was self-administered by nasal spray. The spray was administered to participants 4 times, and each administration consisted of one inhalation, of 4 IU into each nostril. The order of treatments (OT versus placebo) was counterbalanced across participants. The EEG recording began 45 min after the spray administration. Each face, subtending a visual angle of $3.8^\circ \times 4.7^\circ$, was displayed in the center of a gray background for 200 ms. The interstimulus intervals consisted of a fixation cross, with a random duration varying between 800 and 1400 ms. We used different faces from 8 Asian and 8 Caucasian models in the two EEG sessions. Each photo was presented 4 times in a random order in each block, and 4 blocks of 128 trials were included in each EEG session. Participants were informed that they would be shown Asian and Caucasian faces with pain or neutral expressions and that they had to determine the racial identity of each face (Caucasian versus Asian) with a button press using the left or right index finger. The recording procedure in each EEG lasted 20 min.

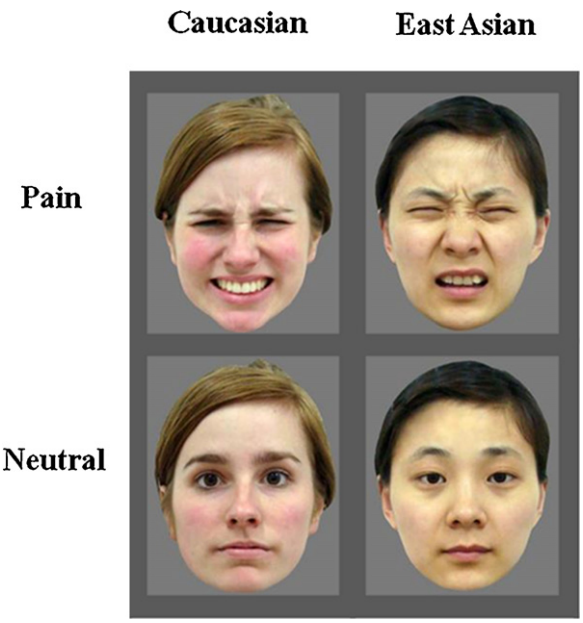

Fig. 1. Illustration of stimuli used in our study.

After the EEG recording, participants rated the intensity of the pain portrayed by each face and their subjective feelings of unpleasantness induced by each face on a 9-point Likert scale (1 = not at all painful or unpleasant, 9 = extremely painful or unpleasant). Participants also rated how much they liked each face on a 9-point Likert scale (1 = not at all, 9 = extremely strong), as an index of their explicit preference for each face. After the second EEG session, each participant completed the IRI (Davis, 1983) to measure his or her empathy trait. After each EEG session, participants were asked to complete a race version of the Implicit Association Test (Greenwald et al., 1998). They categorized Asian faces/positive words with one key and Caucasian faces/negative words with another key in two blocks and Asian faces/negative words with one key and Caucasian faces/positive words with another key in another two blocks. A D score, calculated based on an established algorithm of response latencies (Greenwald et al., 2003), provided an index of participants' implicit attitudes toward racial in-group and out-group faces. A D score larger than zero indicates that in-group faces are associated with a positive rather than negative attitude compared to out-group faces, whereas a D score smaller than zero suggests a negative rather than positive attitude toward in-group faces compared to out-group faces.

2.3. EEG recording and analysis

The EEG referenced to the average of the left and right mastoid electrodes was continuously recorded from 62 scalp electrodes. Eye blinks and vertical eye movements were monitored with electrodes located above and below the left eye. The horizontal electrooculogram was recorded from electrodes placed 1.5 cm lateral to the left and right external canthi. The EEG was amplified (band pass 0.1–100 Hz) and digitized at a sampling rate of 250 Hz. The ERPs in each condition were averaged separately off-line, with an epoch beginning 200 ms before stimulus onset and continuing for 1200 ms. Trials contaminated by eye movements and muscle potentials exceeding $\pm 50\,\mu V$ at any electrode or response errors were excluded from average. This resulted in rejection of $16.9 \pm 10.3\%$ of the trials. The baseline for the ERP measurements was the mean voltage of a 200 ms pre-stimulus interval, and the latency was measured relative to the stimulus onset. The mean amplitudes of each ERP component were calculated at the frontal (Fz, FCz, F3, F4, FC3, and FC4), central (Cz, C3, and C4), parietal (Pz, P3, and P4) and occipito-temporal (P7, P8, PO7, and PO8) electrodes. The analysis of the P2 and N2 components was conducted over the frontal and central electrodes. The parietal and occipital electrodes were included for the analysis of the long latency component, such as the P3, and the early posterior ERP components, such as the N170, respectively. Preliminary repeated measures analyses of variance (ANOVAs) of behavior and ERP data included treatment order (receiving OT or placebo first) as a between-subjects variable. Neither the main effect of treatment order nor its interaction with other variables was significant ($F < 1$). Thus we reported the results of the ANOVAs of reaction times (RTs), response accuracies, and the mean ERP amplitudes with treatment (OT versus placebo), expression (pain versus neutral), and race (Asian versus Caucasian) as within-subjects variables. The ANOVAs of the mean ERP amplitudes recorded at the bilateral electrodes included Hemisphere (electrode over the left versus right hemispheres) as a within-subjects variable. The effect of Hemisphere was not significant ($p > 0.05$) and is not reported in the results section. Voltage topography and standardized Low Resolution

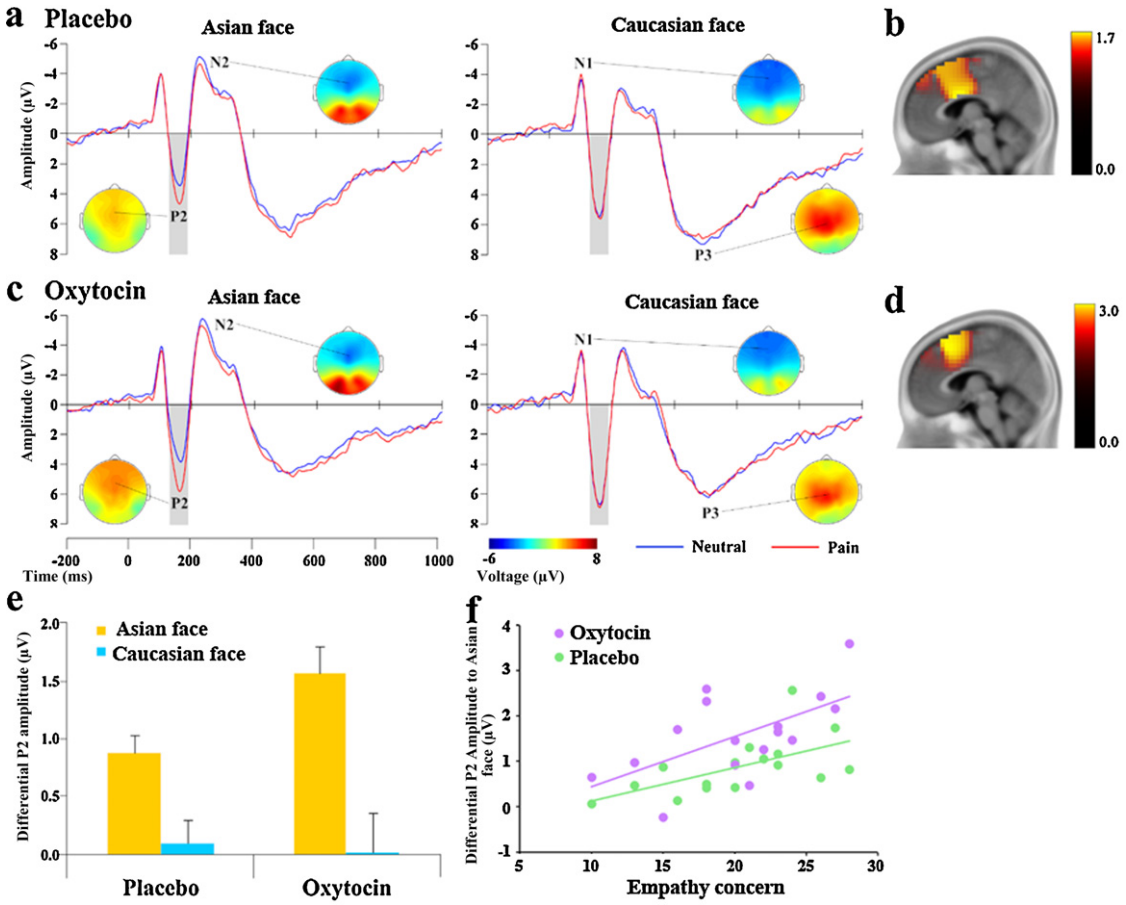

Fig. 2. Illustration of the OT effects on empathic neural responses. (a) ERPs recorded at FCz to pain and neutral expressions after placebo treatment. (b) Source estimation of the neural activity in the P2 time window that differentiated between pain and neutral expressions of Asian faces in the placebo condition. The scale bar represents the log of *F*-ratio for comparisons between ERPs to pain and neutral expressions in the P2 time window. (c) ERPs recorded at FCz to pain and neutral expressions after OT treatment. (d) Source estimation of the neural activity in the P2 time window that differentiated between pain and neutral expressions of Asian faces in the OT condition. (e) The amplitude of the difference wave at 128–188 ms obtained by subtracting ERPs to neutral expression from those to pain expression in the OT and placebo conditions. (f) The correlation between the differential P2 amplitude to pain versus neutral expressions and rating scores of empathy concern in the OT and placebo conditions.

Brain Electromagnetic Tomography (sLORETA, Pascual-Marqui, 2002) were used to estimate potential sources of empathic neural responses.

3. Results

3.1. Behavioral results

The ANOVAs of RTs and accuracy showed significant interactions of race × expression ($F(1,15) = 38.00$ and 14.66, both $p < 0.005$). Participants responded slower ($F(1,15) = 15.27$, $p < 0.001$) and less accurately ($F(1,15) = 5.45$, $p < 0.05$) to pain versus neutral expressions of Asian faces, whereas a reverse pattern was observed for Caucasian faces (RTs, ($F(1,15) = 20.02$, $p < 0.001$; accuracy, $F(1,15) = 10.91$, $p < 0.01$, Table 1). Pain intensity and self-unpleasantness were rated higher on pain expressions than on neutral expressions ($F(1,15) = 168.16$ and 56.11, both $p < 0.001$), but these effects were not modulated by race or treatment (both $p > 0.1$). The explicit likability rating showed a preference for neutral over pain expressions ($F(1,16) = 48.47$, $p < 0.001$). The *D* score in the Implicit Association Test did not differ significantly from zero in the placebo condition ($M = 0.16$, SD = 0.49, $t(1,15) = 1.43$, $p = 0.173$), but was significantly larger than zero in the OT condition ($M = 0.33$, SD = 0.58, $t(15) = 2.36$, $p < 0.05$). These results suggest that, relative to Caucasian faces, Asian faces were significantly associated with a positive rather than negative attitude after OT treatment.

3.2. Electrophysiological results

Fig. 2 illustrates the ERPs at a fronto-central electrode to pain and neutral expressions in the OT and placebo conditions. The ERPs were characterized by a negative wave at 84–116 ms (N1) and a positive deflection at 128–188 ms (P2) over the frontal–central area, which were followed by a negative wave at 200–300 ms (N2) over the frontal region and a long-latency positivity at 400–700 ms (P3) over the parietal area. The face stimuli also elicited a posterior P1 at 88–148 ms and a N170 at 140–180 ms over the occipito-temporal electrodes.

The ANOVAs of the P2 amplitudes at 128–188 ms showed significant main effects of race (Fz: $F(1,15) = 24.81$, $p < 0.001$; FCz: $F(1,15) = 27.14$, $p < 0.001$; Cz: $F(1,15) = 28.44$, $p < 0.001$; F3–F4: $F(1,15) = 20.47$, $p < 0.001$; FC3–FC4: $F(1,15) = 25.89$, $p < 0.001$; C3–C4: $F(1,15) = 33.31$, $p < 0.001$) and expression (Fz: $F(1,15) = 10.94$, $p = 0.005$; FCz: $F(1,15) = 28.07$, $p < 0.001$; Cz: $F(1,15) = 35.19$, $p < 0.001$; F3–F4: $F(1,15) = 8.88$, $p = 0.009$; FC3–FC4: $F(1,15) = 22.23$, $p < 0.001$; C3–C4: $F(1,15) = 24.39$, $p < 0.001$). The P2 amplitudes were increased in response to Caucasian versus Asian faces and in response to expressions of pain versus neutral expressions. These main effects are consistent with previous findings (Ito and Bartholow, 2009; Sheng and Han, 2012) and suggest that the P2 is engaged in coding both race and pain expression. There was a significant main effect of treatment on the P2 amplitude (Fz: $F(1,15) = 7.71$, $p = 0.014$; FCz: $F(1,15) = 5.30$, $p = 0.036$;

Table 1
Behavioral performances and subjective rating scores (mean ± SD).

	Expression	Placebo		Oxytocin	
		Asian	Caucasian	Asian	Caucasian
Reaction time (ms)	Neutral	535 ± 71	533 ± 72	535 ± 57	523 ± 57
	Pain	546 ± 76	522 ± 73	544 ± 70	516 ± 49
Accuracy (%)	Neutral	91 ± 5	90 ± 5	93 ± 5	92 ± 5
	Pain	89 ± 7	92 ± 5	91 ± 7	94 ± 4
Pain intensity	Neutral	2.03 ± 1.22	1.92 ± 1.11	1.82 ± 0.96	1.91 ± 1.04
	Pain	6.80 ± 1.16	6.55 ± 1.40	6.75 ± 1.10	6.65 ± 1.25
Self-unpleasantness	Neutral	2.88 ± 1.65	2.80 ± 1.84	3.01 ± 1.69	2.50 ± 1.34
	Pain	5.48 ± 1.69	5.33 ± 1.35	5.41 ± 1.44	5.66 ± 1.80
Likability	Neutral	4.98 ± 1.02	5.20 ± 1.11	5.02 ± 1.25	5.38 ± 0.96
	Pain	4.10 ± 0.84	4.08 ± 0.91	4.31 ± 1.22	4.37 ± 1.04

F3–F4: $F(1,15) = 6.62$, $p = 0.021$; FC3–FC4: $F(1,15) = 4.95$, $p = 0.042$), as the OT treatment significantly increased the P2 amplitude, compared to the placebo treatment. There was also a significant interaction of expression × race (Fz: $F(1,15) = 19.34$, $p = 0.001$; FCz: $F(1,15) = 22.75$, $p < 0.001$; Cz: $F(1,15) = 29.86$, $p < 0.001$; F3–F4: $F(1,15) = 11.90$, $p = 0.004$; FC3–FC4: $F(1,15) = 17.97$, $p = 0.001$; C3–C4: $F(1,15) = 18.56$, $p = 0.001$), indicating that the effect of pain expression on the P2 amplitude was stronger for Asian than Caucasian faces. Simple effect analysis revealed that the P2 amplitude was enlarged to pain expression, compared to neutral expression of Asian faces (Fz: $F(1,15) = 43.91$, $p < 0.001$; FCz: $F(1,15) = 76.57$, $p < 0.001$; Cz: $F(1,15) = 84.95$, $p < 0.001$; F3–F4: $F(1,15) = 27.59$, $p < 0.001$; FC3–FC4: $F(1,15) = 61.82$, $p < 0.001$; C3–C4: $F(1,15) = 60.68$, $p < 0.001$), but not of Caucasian faces (all $F < 1$). This result replicates our previous findings (Sheng and Han, 2012) and suggests an in-group bias in neural responses to others' suffering. Source estimation suggested that the neural activity in the P2 time window that differentiated between pain and neutral expressions of Asian faces had potential sources in the dorsal ACC and supplementary motor cortex (peak MNI coordinates: 5, 10, 25 and 5, 25, 55 in placebo and OT conditions, respectively; Fig. 2b and d). This result is similar to our previous finding (Sheng and Han, 2012).

Most importantly, the ANOVAs of the P2 amplitudes showed a significant three-way interaction of treatment × race × expression (Fz: $F(1,15) = 5.21$, $p = 0.037$; FCz: $F(1,15) = 4.87$, $p = 0.043$; FC3–FC4: $F(1,15) = 3.19$, $p = 0.094$). Separate analysis revealed that treatment × expression interaction was significant for Asian faces (Fz: $F(1,15) = 3.77$, $p = 0.071$; FCz: $F(1,15) = 6.11$, $p = 0.026$; Cz: $F(1,15) = 5.78$, $p = 0.03$; F3–F4: $F(1,15) = 10.36$, $p = 0.006$; FC3–FC4: $F(1,15) = 4.84$, $p = 0.044$; C3–C4: $F(1,15) = 6.22$, $p = 0.025$), as OT compared to placebo significantly increased the P2 amplitude to pain expressions (Fz: $F(1,15) = 7.63$, $p = 0.015$; FCz: $F(1,15) = 5.87$, $p = 0.028$; F3–F4: $F(1,15) = 6.56$, $p = 0.022$; FC3–FC4: $F(1,15) = 3.89$, $p = 0.069$), but did not affect the P2 amplitude to neutral expressions (all $p > 0.1$). However, the interaction of treatment × expression was not significant for Caucasian faces (all $p > 0.1$), suggesting that OT treatment failed to modulate the P2 amplitude to pain versus neutral expressions of racial out-group members. Separate analysis revealed significant interactions of race × expression in both the placebo (Fz: $F(1,15) = 4.46$, $p = 0.052$; FCz: $F(1,15) = 12.20$, $p = 0.003$; Cz: $F(1,15) = 15.02$, $p = 0.001$; F3–F4: $F(1,15) = 7.204$, $p = 0.017$; FC3–FC4: $F(1,15) = 11.22$, $p = 0.004$; C3–C4: $F(1,15) = 10.35$, $p = 0.006$) and OT conditions (Fz: $F(1,15) = 15.57$, $p = 0.001$; FCz: $F(1,15) = 18.42$, $p = 0.001$; Cz: $F(1,15) = 17.21$, $p = 0.001$; F3–F4: $F(1,15) = 8.21$, $p = 0.012$; FC3–FC4: $F(1,15) = 12.74$, $p = 0.003$; C3–C4: $F(1,15) = 10.49$, $p = 0.006$), suggesting the existence of racial bias in empathic neural responses after placebo and OT treatments.

The differential P2 amplitude to pain versus neutral expressions of Asian faces at the fronto-central electrodes was positively correlated with empathic concern scores across all participants in both the placebo (Fz: $r(16) = 0.608$, $p = 0.012$; FCz: $r(16) = 0.593$, $p = 0.016$; FC3: $r(16) = 0.485$, $p = 0.057$; FC4: $r(16) = 0.701$, $p = 0.003$; C4: $r(16) = 0.523$, $p = 0.038$) and OT conditions (Fz: $r(16) = 0.565$, $p = 0.023$; FCz: $r(16) = 0.600$, $p = 0.014$; Cz: $r(16) = 0.514$, $p = 0.042$; F3: $r(16) = 0.476$, $p = 0.062$; F4: $r(16) = 0.762$, $p = 0.001$; FC3: $r(16) = 0.459$, $p = 0.073$; FC4: $r(16) = 0.620$, $p = 0.010$; C3: $r(16) = 0.431$, $p = 0.095$; C3: $r(16) = 0.615$, $p = 0.011$), suggesting an association between participants' empathy traits and empathic neural responses to others' suffering. To test whether individuals' empathy traits would predict the racial bias in the empathic neural response in the P2 time window, we calculated the correlation between the empathy concern scores and the racial bias in the P2 amplitudes in the placebo and OT conditions, respectively. We found that the empathy concern score positively predicted the racial bias in the empathic neural response in the P2 time window in both the placebo (FCz: $r(16) = 0.560$, $p = 0.024$; FC4: $r(16) = 0.571$, $p = 0.021$; Cz: $r(16) = 0.493$, $p = 0.052$; C3: $r(16) = 0.516$, $p = 0.041$; C4: $r(16) = 0.578$, $p = 0.019$) and OT conditions (Fz: $r(16) = 0.516$, $p = 0.041$; F4: $r(16) = 0.541$, $p = 0.030$; FCz: $r(16) = 0.507$, $p = 0.045$). We also examined whether individuals' empathy traits were associated with the effect of OT on the racial bias in the P2 amplitudes, which was defined by subtracting the racial bias in the P2 amplitudes (i.e., the difference in the P2 amplitude to pain versus neutral expressions between racial in-group and out-group members) in the placebo condition from that in the OT condition. However, there was no significant correlation between the empathy concern score and the effect of OT on the racial bias in the empathic neural response in the P2 time window (all $p > 0.1$).

Next we assessed whether the racial bias in the empathic neural responses was associated with the racial bias in the implicit racial attitudes. The *D* score in the Implicit Association Test was used as a behavioral index of the racial bias in implicit attitudes. The difference in the empathic neural responses between Asian and Caucasian faces, which was calculated by subtracting the differential P2 amplitudes to pain versus neutral expressions of Caucasian faces from those of Asian faces, was used as an index of the racial bias in empathic neural responses. We found that the racial bias in the empathic neural responses positively correlated with the *D* score in the OT condition (FCz: $r(16) = 0.551$, $p = 0.027$; Cz: $r(16) = 0.602$, $p = 0.014$; F3: $r(16) = 0.443$, $p = 0.085$; F4: $r(16) = 0.505$, $p = 0.046$; FC3: $r(16) = 0.453$, $p = 0.078$; FC4: $r(16) = 0.536$, $p = 0.032$; C3: $r(16) = 0.433$, $p = 0.094$; C3: $r(16) = 0.572$, $p = 0.021$), but not in the placebo condition ($p > 0.3$, Fig. 3), suggesting an enhancement of the association between the implicit positive attitude toward racial

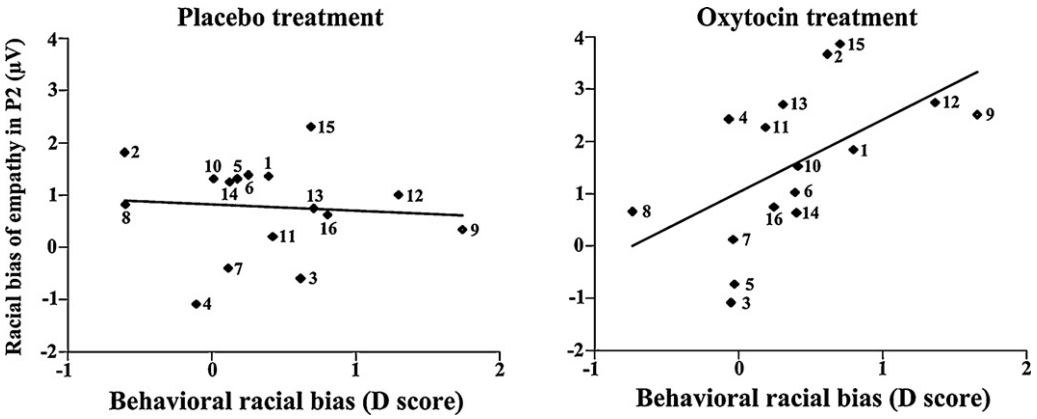

Fig. 3. Illustration of the correlation between the racial bias in empathic neural responses and the *D* score in the placebo and OT conditions, respectively. Each individual participant was indicated with a number.

in-group members and the racial bias in empathic neural responses in the P2 time window, after the OT treatment.

The ANOVAs of the N2 amplitudes showed significant main effects of race (Fz: $F(1,15) = 49.35$, $p < 0.001$; FCz: $F(1,15) = 47.61$, $p < 0.001$; Cz: $F(1,15) = 49.65$, $p < 0.001$; F3–F4: $F(1,15) = 36.28$, $p < 0.001$; FC3–FC4: $F(1,15) = 46.24$, $p < 0.001$; C3–C4: $F(1,15) = 53.36$, $p < 0.001$) and expression (Fz: $F(1,15) = 3.49$, $p = 0.081$; FCz: $F(1,15) = 7.87$, $p = 0.013$; Cz: $F(1,15) = 6.74$, $p = 0.020$; FC3–FC4: $F(1,15) = 5.56$, $p = 0.032$; C3–C4: $F(1,15) = 3.91$, $p = 0.067$), due to that the N2 was of larger amplitude to Asian than Caucasian faces and to neutral than pain expressions (Fig. 2a and c). There were also significant main effects of race on P3 amplitude (Pz: $F(1,15) = 8.38$, $p = 0.011$; P3–P4: $F(1,15) = 6.29$, $p = 0.023$) and N170 amplitudes (P7–P8: $F(1,15) = 38.13$, $p < 0.001$; PO7–PO8: $F(1,15) = 26.01$, $p < 0.001$), suggesting larger P3 for Caucasian faces than for Asian faces and larger N170 amplitudes for Asian faces than for Caucasian faces. The ANOVAs of the N2, P3, P1, and N170 amplitudes showed that neither the main effect of treatment nor its interaction with race and expression was significant (all $p > 0.1$). Correlation analyses failed to find a significant correlation between the *D* score in the Implicit Association Test and the race effect on the P2, N170, N2, and P3 components (all $p > 0.1$).

4. Discussion

The modulation of the P2 amplitude by facial expression of pain is consistent with the previous findings that perception of human body parts (e.g., hand or foot) receiving painful versus neutral stimulation elicits increased positivity over the fronto-central region (Fan and Han, 2008; Han et al., 2008; Li and Han, 2010; Decety et al., 2010). The source estimation suggests that the P2 empathic neural responses might arise from the midcingulate and the supplementary motor area. Moreover, the P2 empathic response was greater to racial in-group faces than to out-group faces. The P2 effect is consistent with the previous findings of a racial in-group bias in empathic neural responses within the same time window (Sheng and Han, 2012) and in a similar brain region (Xu et al., 2009). Moreover, we found that, relative to the placebo treatment, the OT treatment selectively increased neural responses to pain expression of racial in-group faces in the context of racial categorization and thus increased the racial bias in empathic neural responses in the P2 time window.

Although behavioral research suggests that OT facilitates understanding or sharing of others' emotions (Domes et al., 2007; Hurlemann et al., 2010; Bartz et al., 2010), there has been no evidence for the modulation of empathic neural responses by OT treatment. Singer et al. (2008) found that, relative to treatment with a placebo, OT treatment reduced amygdala activation when participants received painful stimulation themselves but did not modulate empathy-relevant brain activation in the anterior insula. This study did not investigate the OT effects on empathic neural responses in a specific social context. Our ERP findings suggest an effect of OT that was specific to an in-group versus out-group context and support the existence of an interaction between social (e.g., intergroup relationship) and biological (e.g., OT) factors in the modulation of empathic neural responses to perceived pain in others.

The effect of OT on empathic neural responses took place between 100 and 200 ms after sensory stimulation. Sheng and Han (2012) showed that empathic neural responses in this time window were modulated by manipulation of cognitive strategies and intergroup relationships. Enhanced attention to an individual's feelings and inclusion of other-race individuals on one's own team for competitions reduced the racial bias in empathic neural responses, by increasing empathic neural activity to other-race individuals rather than by decreasing empathic neural activity to same-race individuals. Unlike the manipulation of cognitive strategies and intergroup relationships, intranasally administered OT increased the empathic neural responses in the P2 time window to same-race individuals but produced little effect on the P2 empathic neural responses to other-race individuals. Thus P2 empathic neural responses to same-race and other-race individuals seem to be sensitive to psychological manipulations and neuropeptide, respectively.

Interestingly, neither intranasally administered OT (the current work) nor manipulation of cognitive strategies and intergroup relationships (Sheng and Han, 2012) affected the rating scores of self-reported unpleasantness induced by viewing pain expressions. Rating scores are explicit measurements of subjective feelings and are sensitive to social contexts and social desire. It is likely that our participants were concerned about overtly expressing greater empathy for racial in-group members than for out-group members because racial in-group bias is apparently not encouraged by current societies. OT treatment appeared to modulate participants' implicit attitudes toward racial in-group members because the *D* score of the Implicit Associate Test was larger than zero after the OT treatment. The OT treatment resulted in a significant association between racial bias in empathic neural responses and participants' implicit attitudes toward racial in-group faces. The previous studies have shown that OT treatment significantly affects attitudes, such as social trust, toward others (Kosfeld et al., 2005; Baumgartner et al., 2008; De Dreu, 2012). One possibility is that, in our study, the OT treatment might have changed participants' implicit attitudes toward racial in-group and out-group members. The resulting

sustained variation of implicit attitudes might have modified the neural activity to perceived pain in racial in-group members in a top-down manner. This possibility should be investigated in future research.

Empathic neural responses are associated with altruistic behavior. Neural activity to perceived pain predicts how much money participants donate (Ma et al., 2011) and how often participants sacrifice themselves to help in-group members (Hein et al., 2010). The racial bias in empathy-related neural activity may lead to racial in-group preference during altruistic behavior. Indeed, individuals with racial bias in empathy tend to assign more lenient punishments (Johnson et al., 2002) and show pain treatment biases (Drwecki et al., 2011) toward racial in-group compared to out-group members. The racial bias in empathy may reflect an evolutionary strategy to prevent an inappropriate extension of in-group generosity to out-group members in order to benefit the survival of in-group individuals. Recent research suggests that OT motivates in-group favoritism and parochial cooperation instead of creating more benevolent views of others generally (De Dreu et al., 2010, 2011). The effect of OT on empathic neural responses to in-group members may play a role in the modulation of social behavior toward in-group members. OT is a highly reserved neuropeptide and has been a hormone throughout evolution (Meyer-Lindenberg et al., 2011). It may signal its adaptive value in protecting in-group benefit by facilitating in-group favoritism in empathy.

There has been evidence for OT engagement in first-hand pain experience. Animal studies have shown that OT administration reduces pain sensitivity to thermal heat (Agren et al., 1997) and mechanical pain (Petersson et al., 2001). In contrast, an OT antagonist increases pain sensitivity (Uvnas-Moberg et al., 1992). Intrathecal OT administration in humans reduces pain in individuals with acute or chronic low back pain (Yang et al., 2002). The effect of OT on pain experience may arise from an enhancement of endogenous opioid activity (Miranda-Cardenas et al., 2006) and a reduction of sympathetic nervous system activity (Sofroniew, 1980). These studies did not consider whether the effect of OT on pain experience is influenced by the social relationship between a giver and a receiver of pain stimulation. Our findings indicate that the effect of OT on empathic neural responses to others' suffering is sensitive to the social relationship between an observer and a target person. Future research should clarify whether the effect of OT characterizes the key difference between first-hand pain experience and empathy for others' pain.

A recent behavioral study showed that, relative to placebo treatment, OT treatment increased the feeling of envy when an individual gained less money than another player and increased the feeling of gloating when one player gained more money than the other (Shamay-Tsoory et al., 2009). Thus, OT may enhance the social comparison that produce a negative effect on prosocial behaviors. Hein et al. (2010) found that, when seeing out-group members in pain, participants with more negative impressions of out-group members showed stronger activity in the right nucleus accumbens, a brain region that has been associated with schadenfreude (Takahashi et al., 2009). These findings leave an open question of whether OT influences the neural activity in the reward-related system while perceiving out-group members' pain.

Previous studies have shown that other facial expressions also modulate the P2 amplitude. Kubota and Ito, 2007 recorded ERP to black and white faces from Caucasians. They found enlarged P2 amplitudes to angry and happy faces compared to neutral faces. The P2 modulation by angry/happy expressions did not differ between racial in-group and out-group faces. The P2 modulation by pain expression seemed to be different from that observed by Kubota and Ito, 2007, in terms of the effect of the racial relationship between an observer and a target person. Previous research has shown that the P2 is sensitive to novel or negative stimuli because the P2 is enlarged by negative-arousing pictures (Bar-Haim et al., 2005) and threat-related pictures or words (Taake et al., 2009; Thomas et al., 2007; Weymar et al., in press). These findings suggest that the P2 amplitude may reflect enhanced attention to novel stimuli that are relevant to one's own safety. From an evolutionary perspective, intergroup competition for resources results in antagonism between in-group and out-group members and leads to the stereotype that out-group members are dangerous. Faces of out-group members may be perceived with higher novelty compared to faces of in-group members regardless of facial expression (e.g., painful versus neutral faces). This hypothesis may explain the race effect on the P2 amplitude observed in our study and others. In-group members are usually not dangerous. Pain expression of an in-group member may signal a need for help and have higher novelty compared to neutral faces. This impression may result in higher sensitivity to pain expression of in-group versus out-group members, as indicated by the greater P2 amplitude and delayed responses to pain versus neutral expressions of racial in-group members in a race judgment task. The P2 modulation by in-group members' pain expressions in particular is associated with empathy because we showed that the P2 amplitude to pain versus neutral expressions was correlated with an individual's empathy concern capacity.

In conclusion, our ERP results showed that, relative to the placebo treatment, the OT treatment increased the empathic neural responses to racial in-group faces at 128–188 ms after stimulus onset but failed to modulate the empathic neural responses to racial out-group faces. Our findings suggest that OT may interact with the social relationship between an observer and a target person to modulate human empathy for the suffering of others. Future research should address how the interaction between the social relationship and biological factors, such as OT, influences human social behaviors.

Acknowledgements

This work was supported by Beijing Municipal Natural Science Foundation (No. Z111107067311058), the National Natural Science Foundation of China (Project 30910103901, 91024032, 81161120539), and the National Basic Research Program of China (973 Program 2010CB833903). We thank Kate Woodcock for proofreading the manuscript.

References

Agren, G., Uvnas-Moberg, K., Lundeberg, T., 1997. Olfactory cues from an oxytocin injected male rat can induce anti-nociception in its cagemates. Neuroreport 8, 3073–3076.

Avenanti, A., Bueti, D., Galati, G., Aglioti, S.M., 2005. Transcranial magnetic stimulation highlights the sensorimotor side of empathy for pain. Nature Neuroscience 8, 955–960.

Avenanti, A., Sirigu, A., Aglioti, S., 2010. Racial bias reduces empathic sensorimotor resonance with other-race pain. Current Biology 20, 1018–1022.

Bar-Haim, Y., Lamy, D., Glickman, S., 2005. Attentional bias in anxiety: a behavioral and ERP study. Brain and Cognition 59, 11–22.

Bartz, J.A., Zaki, J., Bolger, N., Hollander, E., Ludwig, N.N., Kolevzon, A., Ochsner, K.N., 2010. Oxytocin selectively improves empathic accuracy. Psychological Science 21, 1426–1428.

Bartz, J.A., Zaki, J., Bolger, N., Ochsner, K.N., 2011. Social effects of oxytocin in humans: context and person matter. Trends in Cognitive Sciences 15, 301–309.

Batson, C.D., 1998. Altruism and prosocial behavior. In: Gilbert, D.T., Fisk, S., Gardner, L. (Eds.), The Handbook of Social Psychology. McGraw Hill, Boston, pp. 282–316.

Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U., Fehr, E., 2008. Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. Neuron 58, 639–650.

Davis, M.H., 1983. The effects of dispositional empathy on emotional reactions and helping: a multidimensional approach. Journal of Personality 51, 167–184.

De Dreu, C.K.W., Greer, L.L., Handgraaf, M.J.J., Shalvi, S., Van Kleef, G.A., Baas, M., Ten Velden, F.S., Van Dijk, E., Feith, S.W., 2010. The neuropeptide oxytocin regulates parochial altruism in intergroup conflict among humans. Science 328, 1408–1411.

De Dreu, C.K.W., Greer, L.L., Van Kleef, G.A., Shalvi, S., Handgraaf, M.J.J., 2011. Oxytocin promotes human ethnocentrism. Proceedings of the National Academy of Sciences of the United States of America 108, 1262–2126.
De Dreu, C.K.W., 2012. Oxytocin modulates cooperation within and competition between groups: an integrative review and research agenda. Hormones and Behavior 61, 419–428.
Decety, J., Yang, C., Cheng, Y., 2010. Physicians down-regulate their pain empathy response: an event-related brain potential study. NeuroImage 50, 1676–1682.
de Waal, F.B., 2008. Putting the altruism back into altruism: the evolution of empathy. Annual Review of Psychology 59, 279–300.
Domes, G., Heinrichs, M., Michel, A., Berger, C., Herpertz, S., 2007. Oxytocin improves mind-reading in humans. Biological Psychiatry 61, 731–733.
Drwecki, B.B., Moore, C.F., Ward, S.E., Prkachin, K.M., 2011. Reducing racial disparities in pain treatment: the role of empathy and perspective-taking. Pain 152, 1001–1006.
Fan, Y., Han, S., 2008. Temporal dynamic of neural mechanisms involved in empathy for pain: an event-related brain potential study. Neuropsychologia 46, 160–173.
Greenwald, A., McGhee, D., Schwartz, J., 1998. Measuring individual differences in implicit cognition: the implicit association test. Journal of Personality and Social Psychology 74, 1464–1480.
Greenwald, A., Nosek, B., Banaji, M., 2003. Understanding and using the Implicit Association Test: I. An improved scoring algorithm. Journal of Personality and Social Psychology 85, 197–216.
Gu, X., Han, S., 2007. Attention and reality constraints on the neutral processes of empathy for pain. NeuroImage 36, 256–267.
Han, S., Fan, Y., Mao, L., 2008. Gender difference in empathy for pain: an electrophysiological investigation. Brain Research 1192, 85–93.
Han, S., Fan, Y., Xu, X., Qin, J., Wu, B., Wang, X., Aglioti, S.M., Mao, L., 2009. Empathic neural responses to others' pain are modulated by emotional contexts. Human Brain Mapping 30, 3227–3237.
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, 149–160.
Hurlemann, R., Patin, A., Onur, O., Cohen, M., Baumgartner, T., Metzler, S., Dziobek, I., Gallinat, J., Wagner, M., Maier, W., Kendrick, K.M., 2010. Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. Journal of Neuroscience 30, 4999–5007.
Ito, T.A., Bartholow, B.D., 2009. The neural correlates of race. Trends in Cognitive Sciences 13, 524–531.
Johnson, J.D., Simmons, C.H., Jordan, A., MacLean, L., Taddei, J., Thomas, D., 2002. Rodney King and O.J. revisited: the impact of race and defendant empathy induction on judicial decisions. Journal of Applied Social Psychology 32, 1208–1223.
Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U., Fehr, E., 2005. Oxytocin increases trust in humans. Nature 435, 673–676.
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, 738–752.
Li, W., Han, S., 2010. Perspective taking modulates event related potentials to perceived pain. Neuroscience Letters 469, 328–332.
Ma, Y., Wang, C., Han, S., 2011. Neural responses to perceived pain in others predict real-life monetary donations in different socioeconomic contexts. NeuroImage 57, 1273–1280.
Meyer-Lindenberg, A., Domes, G., Kirsch, P., Heinrichs, M., 2011. Oxytocin and vasopressin in the human brain: social neuropeptides for translational medicine. Nature Review Neuroscience 12, 524–538.
Miranda-Cardenas, Y., Rojas-Piloni, G., Martinez-Lorenzana, G., Rodriguez-Jimenez, J., Lopez-Hidalgo, M., Freund-Mercier, M., Condes-Lara, M., 2006. Oxytocin and electrical stimulation of the paraventricular hypothalamic nucleus produce antinociceptive effects that are reversed by an oxytocin antagonist. Pain 122, 182–189.
Pascual-Marqui, R.D., 2002. Standardized low-resolution brain electromagnetic tomography (sLORETA): methods and findings. Experimental and Clinical Pharmacology 24 (Suppl D), 5–12.
Petersson, M., Wiberg, U., Lundeberg, T., Uvnas-Moberg, K., 2001. Oxytocin decreases carrageenan induced inflammation in rats. Peptides 22, 1479–1484.
Rainville, P., Duncan, G.H., Price, D.D., Carrier, B., Bushnell, M.C., 1997. Pain affect encoded in human anterior cingulate but not somatosensory cortex. Science 277, 968–971.
Saarela, M.V., Hlushchuk, Y., Williams, A.C., Schurmann, M., Kalso, E., Hari, R., 2007. The compassionate brain: humans detect intensity of pain from another's face. Cerebral Cortex 17, 230–237.
Shamay-Tsoory, S.G., Fischer, M., Dvash, J., Harari, H., Perach-Bloom, N., Levkovitz, Y., 2009. Intranasal administration of oxytocin increases envy and schadenfreude (gloating). Biological Psychiatry 66, 864–870.
Sheng, F., Han, S., 2012. Manipulations of cognitive strategies and intergroup relationships reduce the racial bias in empathic neural responses. NeuroImage 61, 786–797.
Singe, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R., Frith, C., 2004. Empathy for pain involves the affective but not sensory components of pain. Science 303, 1157–1162.
Singer, T., Snozzi, R., Bird, G., Petrovic, P., Silani, G., Heinrichs, M., Dolan, R.J., 2008. Effects of oxytocin and prosocial behavior on brain responses to direct and vicariously experienced pain. Emotion 8, 781–791.
Sofroniew, M.V., 1980. Projections from vasopressin, oxytocin, and neurophysin neurons to neural targets in the rat and human. Journal of Histochemistry and Cytochemistry 28, 475–478.
Taake, I., Jaspers-Fayer, F., Liotti, M., 2009. Early frontal responses elicited by physical threat words in an emotional Stroop task: modulation by anxiety sensitivity. Biological Psychology 81, 48–57.
Takahashi, H., Kato, M., Matsuura, M., Mobbs, D., Suhara, T., Okubo, Y., 2009. When your gain is my pain and your pain is my gain: neural correlates of envy and schadenfreude. Science 323, 937–939.
Thomas, S.J., Johnstone, S.J., Gonsalvez, C.J., 2007. Event-related potentials during an emotional Stroop task. International Journal of Psychophysiology 63, 221–231.
Uvnas-Moberg, K., Bruzelius, G., Alster, P., Bileviciute, I., Lundeberg, T., 1992. Oxytocin increases and a specific oxytocin antagonist decreases pain threshold in male rats. Acta Physiologica Scandinavica 144, 487–488.
Wager, T.D., Rilling, J.K., Smith, E.E., Sokolik, A., Casey, K.L., Davidson, R.J., Kosslyn, S.M., Rose, R.M., Cohen, J.D., 2004. Placebo-induced changes in FMRI in the anticipation and experience of pain. Science 303, 1162–1167.
Weymar, M., Bradley, M.M., Hamm, A.O., Lang, P.J. When fear forms memories: threat of shock and brain potentials during encoding and recognition. Cortex, in press.
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, 8525–8529.
Yang, Z., Wheatley, M., Coote, J., 2002. Neuropeptides, amines and amino acids as mediators of the sympathetic effects of paraventricular nucleus activation in the rat. Experimental Physiology 87, 663–674.