

Oxytocin Modulates Attention Switching Between Interoceptive Signals and External Social Cues

Shuxia Yao^{1,2}, Benjamin Becker^{1,2}, Weihua Zhao¹, Zhiying Zhao¹, Juan Kou¹, Xiaole Ma¹, Yayuan Geng¹, Peng Ren¹ and Keith M Kendrick^{*,1,3}

¹Key Laboratory for NeuroInformation, School of Life Science and Technology, Center for Information in Medicine, University of Electronic Science and Technology of China, Chengdu, China

Emotional experience involves an integrated interplay between processing of external emotional cues and interoceptive feedback, and this is impaired in a number of emotional disorders. The neuropeptide oxytocin (OT) enhances the salience of external social cues but its influence on interoception is unknown. The present pharmaco-fMRI study therefore investigated whether OT enhances interoceptive awareness and if it influences the interplay between interoceptive and salience processing. In a randomized, double-blind, between-subject, design study 83 subjects received either intranasal OT or placebo. In Experiment 1, subjects performed a heartbeat detection task alone, while in Experiment 2 they did so while viewing both neutral and emotional face stimuli. Interoceptive accuracy and neural responses in interoceptive and salience networks were measured. In Experiment 1, OT had no significant influence on interoceptive accuracy or associated activity in the right anterior insula (AI) and dorsal anterior cingulate cortex. However, in Experiment 2 when face stimuli were also presented, OT decreased interoceptive accuracy and increased right AI activation and its functional connectivity with the left posterior insula (PI), with the latter both being negatively correlated with accuracy scores. The present study provides the first evidence that while OT does not influence processing of interoceptive cues *per se* it may switch attention away from them towards external salient social cues by enhancing right AI responses and its control over the PI. Thus OT may help regulate the interplay between interoceptive and external salience processing within the insula and could be of potential therapeutic benefit for emotional disorders.

Neuropsychopharmacology (2018) **43**, 294–301; doi:10.1038/npp.2017.189; published online 27 September 2017

INTRODUCTION

Interoception, defined as the sense of the physiological condition of the body (Craig, 2002), provides an important link between peripheral physiological changes and brain cognitive and affective processes. It is not only crucial for maintenance of internal homeostatic state but also underlies the subjective experience of emotions, as proposed by the James–Lange theory of emotion and Damasio’s somatic marker hypothesis (Lange and James, 1967; Damasio, 1994). The role of interoception in emotional experience and its dysfunction in psychiatric disorders have been consistently highlighted (Critchley and Harrison, 2013; Seth, 2013). Imaging studies have demonstrated an overlap between core networks involved in emotional and interoceptive

processing, particularly the insula, with the anterior insula (AI) being involved in cognitive and emotional aspects and the posterior insula (PI) being specifically involved in primary sensory components of interoceptive signals, as well as the dorsal anterior cingulate cortex (dACC) (Critchley *et al*, 2004; Pollatos *et al*, 2007b; Kurth *et al*, 2010; Chang *et al*, 2012; Zaki *et al*, 2012; Uddin *et al*, 2014).

The hypothalamic neuropeptide oxytocin (OT) plays an important role in human social cognition and motivation (Bartz *et al*, 2011; Striepens *et al*, 2011). In particular, it has been reported that OT can improve attention towards (Domes *et al*, 2013; Xu *et al*, 2015) and recognition of (Fischer-Shofty *et al*, 2010; Shahrestani *et al*, 2013) emotional faces. Oxytocin can also facilitate emotional memory (Guastella *et al*, 2008) and empathy (Hurlemann *et al*, 2010). However, no study to date has investigated whether OT influences behavioral or neural sensitivity to interoceptive information. There are OT receptors present in a number of visceral organs such as the heart and stomach, as well as in core interoceptive brain regions such as the insula and cingulate cortex (Gimpl and Fahrenholz, 2001; Jankowski *et al*, 2004). Intranasal OT can also increase heart-rate variability (Kemp *et al*, 2012).

On the other hand, an influential hypothesis concerning OT’s functional effects is that it enhances the salience of external social cues (Bartz *et al*, 2011; Gao *et al*, 2016;

*Correspondence: Professor KM Kendrick, Key Laboratory for NeuroInformation of Ministry of Education, Center for Information in Medicine, University of Electronic Science and Technology of China No. 2006, Xiyuan Ave., West Hi-Tech Zone, Chengdu, Sichuan 611731, China, Tel: +86 28 61830811, Fax: +86 28 61830811, E-mail: k.kendrick.uestc@gmail.com

²These authors contributed equally to this work.

³Current address: No. 2006, Xiyuan Ave., West Hi-Tech Zone, Chengdu, Sichuan 611731, China.

Received 25 May 2017; revised 16 August 2017; accepted 20 August 2017

Shamay-Tsoory and Abu-Akel, 2016). The brain salience network also involves the insula and anterior cingulate and this overlap between neural processing of internal interoceptive cues and the salience of external cues therefore permits optimal integration and switching of attention between them. A recent review has hypothesized that OT may facilitate integration of interoceptive and external cues (Quattrocki and Friston, 2014), and suggested that such integration may be dysfunctional in autistic individuals. However, others have proposed instead that this integration may be more pertinent with respect to alexithymia (Brewer et al, 2015). Intranasal OT can also shift the normal self-bias (internal) more towards consideration of the value of others (external) (Abu-Akel et al, 2015; Bartz et al, 2015; Zhao et al, 2016), in line with its well-established role in promoting formation and maintenance of affiliative bonds (Striepens et al, 2011).

The aim of the current study was therefore first to investigate whether OT alters interoceptive awareness and neural processing *per se* and secondly whether in the presence of external social cues it switches attention away from internal interoceptive cues towards external ones. We investigated the modulatory effects of OT on interoceptive accuracy and associated neural responses using a revised heartbeat detection task in conjunction with functional MRI (fMRI) and simultaneous electrocardiography (ECG) (Experiment 1). Subjects then performed the same interoceptive task while they viewed face stimuli (Experiment 2). Given OT's effects on facilitating emotional processing (Fischer-Shofty et al, 2010; Shahrestani et al, 2013) and the close associations between interoception and emotion (Pollatos et al, 2005; Pollatos et al, 2007a), we hypothesized that in Experiment 1 OT would increase interoceptive accuracy and corresponding activity in interoceptive networks. However in Experiment 2, based on the social salience hypothesis of OT (Shamay-Tsoory and Abu-Akel, 2016) and especially findings that OT could improve attention towards emotional and neutral faces (Domes et al, 2013; Xu et al, 2015), we hypothesized that in the presence of salient external facial cues it would have the opposite effect. We further hypothesized that interoceptive accuracy, and the OT's effects on it, would be modulated differentially by specific face emotions.

MATERIALS AND METHODS

Participants and Treatment

In total 83 healthy male students participated in a randomized, double-blind, between-subject experiment where they received either intranasal 40 IU OT or placebo (PLC; details see Supplementary information). To control for potential confounding effects from personality traits or mood states, subjects completed Chinese versions of validated psychometric questionnaires before treatment (Supplementary information). Subjects received treatment using a standard protocol and the task started in the MRI scanner 45 min later (Guastella et al, 2013). In post-experiment interviews subjects were unable to identify better than chance whether they had received OT or PLC. Written informed consent was obtained before study inclusion. All procedures were approved by the local ethical committee and were in accordance with the latest

version of the Declaration of Helsinki. This study was registered as clinical trial (NCT03074409, <https://clinicaltrials.gov/show/NCT03074409>).

Experiment 1

A revised heartbeat detection paradigm was used in Experiment 1 (Zaki et al, 2012), which consisted of two tasks. In the heartbeat detection task participants were instructed to make a keypress each time they felt their heartbeat, while in a dot detection control task participants were instructed to make a keypress each time they detected a white dot appearing on the screen (Supplementary information).

Experiment 2

Experiment 2 investigated whether OT switched attention away from internal interoceptive cues towards external social salient ones by investigating the effects of simultaneous presentation of face stimuli. In this task, subjects were instructed to make a keypress response each time they felt their heartbeat while watching emotional or neutral faces or a scrambled mosaic of neutral faces (Supplementary information).

Image Acquisition and Data Analysis

Data were acquired using a 3T, GE MR750 scanner and an MRI-compatible ECG system. MRI data were processed using standardized procedures in SPM8 (Friston et al, 1994) (details see Supplementary information).

In Experiment 1, the first-level design matrix included 4 regressors (interoception task, control task, interoception task cue, control task cue). Contrast images of "interoception > control" were created for each subject. In Experiment 2, 5 regressors (fearful, disgust, happy, neutral, and scrambled faces) were included in the design matrix. They were convolved with the canonical hemodynamic response function with the six head-motion parameters as confounding variables. Contrast images for each face type, all emotions, and between each emotional and neutral faces were created. On the second (between-subject) level, group differences were analyzed using two-sample *t*-tests. Main effects of conditions were analyzed using one-sample tests. Interaction between face types and treatment was tested using a flexible factorial design. On the basis of previous studies (Critchley et al, 2004; Pollatos et al, 2007b), and to specifically test our hypotheses, analyses focused on a *priori* regions of interests (ROIs) including AI and dACC as key interoceptive processing hubs. As a key OT-sensitive and emotional processing region, amygdala was additionally included as a ROI in Experiment 2. The insula and ACC were anatomically defined using Automated Anatomic Labeling atlas (Tzourio-Mazoyer et al, 2002). Amygdala was derived from probabilistic maps in Anatomy toolbox 2.1 (Eickhoff et al, 2005). Within these a *priori* masks, a threshold of $p < 0.05$ family-wise error (FWE) corrected at peak level was set for multiple comparisons using small volume correction (SVC). For regions other than the a *priori* ROIs, an exploratory whole-brain analysis was performed using $P_{FWE} < 0.05$ peak correction.

To further examine OT's effects during interoceptive processing, we performed exploratory psychophysiological interaction (PPI) analyses using the PPI toolbox (McLaren et al, 2012). Seed regions were defined as 6-mm spheres centered at the maximum t -value of between-group differences from the BOLD-level analysis. Two-sample t -tests were used to examine group connectivity differences. Given the exploratory nature of this analysis an uncorrected $p < 0.001$ threshold was used and only clusters larger than 10 voxels were reported.

RESULTS

Independent t -tests on pre-treatment questionnaire scores measuring autistic traits, empathy, and personality revealed no significant differences between the OT and PLC groups ($p > 0.151$; Supplementary Table S1). Pre- and post-treatment measures of mood also revealed no significant effects of OT ($p > 0.394$; Supplementary information).

Experiment 1

Behavioral results. To examine OT's effects on interoceptive accuracy, a repeated-measures ANOVA was performed on interoceptive accuracy scores with task-condition (heartbeat detection vs control) as within-subject factor and treatment (OT vs PLC) as between-subject factor. This revealed a significant main effect of task ($F(1, 73) = 57.34$, $p < 0.001$), showing that subjects performed better in the control than in the heartbeat detection task (0.71 ± 0.20 vs 0.90 ± 0.07). There was no significant main effect of treatment ($F(1, 73) = 2.08$, $p = 0.153$) or interaction ($F(1, 73) = 1.06$, $p = 0.306$).

fMRI results. To validate our paradigm, interoceptive networks were analyzed in an initial step (interoception > control task). Results showed that the paradigm induced robust activity in the left insula (AI: MNI = -30, 26, -8; $t = 7.03$; $P_{FWE} < 0.001$ SVC; voxels = 959; PI: MNI = -40, -16, 14; $t = 4.50$; $P_{FWE} = 0.004$ SVC; voxels = 27), right insula (right PI extending to AI: MNI = 44, -14, 14; $t = 7.77$ for PI and MNI = 30, 22, -10; $t = 7.46$ for AI; $P_{FWE} < 0.001$ SVC; voxels = 1130; Figure 1a), and bilateral dACC (left: MNI = -6, 30, 30; $t = 7.87$; $P_{FWE} < 0.001$ SVC; voxels = 705; right: MNI = 10, 30, 24; $t = 6.94$; $P_{FWE} < 0.001$ SVC; voxels = 710; Figure 1b). There were, however, no significant modulatory effects of treatment (ie, interactions) in the a priori ROIs ($P_{FWE} < 0.05$ SVC). Regions other than the a priori ROIs using a whole-brain analysis ($P_{FWE} < 0.05$) are reported in Supplementary Table S2.

Experiment 2

Behavioral results. To examine whether OT's effects on interoceptive accuracy varied as a function of the simultaneous presentation of face stimuli, a repeated-measures ANOVA was performed on interoceptive accuracy scores, with face types as within-subject and treatment as between-subject factors. Results showed a significant main effect of face types ($F(4, 292) = 8.46$, $p < 0.001$). Post-hoc Bonferroni-corrected tests revealed that subjects performed the heartbeat

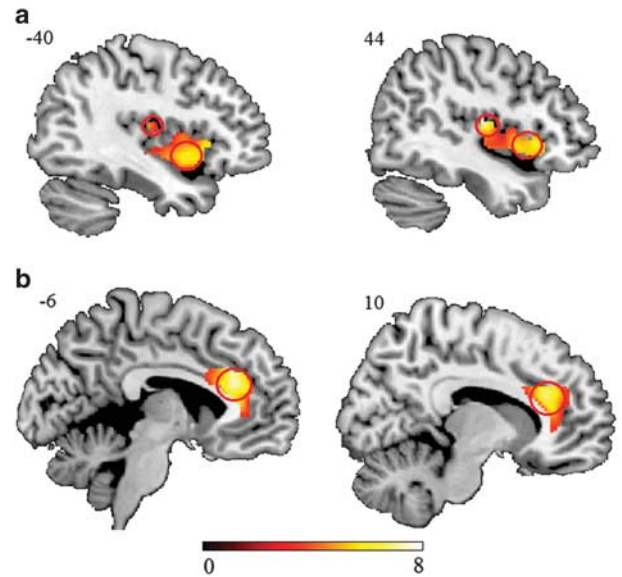


Figure 1 The main effect of task (interoceptive task > control task) on interoceptive network activity in Experiment 1. (a) Bilateral anterior insula and posterior insula. (b) Bilateral dorsal anterior cingulate cortex. Statistical maps are displayed with a threshold of $p < 0.001$ uncorrected.

detection task best when presented with disgust faces and worst with scrambled faces (Figure 2a). The main effect of treatment was also significant ($F(1, 73) = 4.00$, $p = 0.049$), with subjects in the PLC group showing greater accuracy than those in the OT group (Figure 2b), but no significant interaction between face types and treatment ($F(4, 292) = 0.20$, $p = 0.936$).

fMRI results. We first examined the main effect of face types and found robust activity in regions involved in emotional and interoceptive processing, including bilateral AI, dACC, and amygdala ($P_{FWE} < 0.05$; Supplementary Table S3). The main effect of treatment (OT > PLC) across face types was then examined and in contrast to Experiment 1 found stronger right AI activity in the OT compared with the PLC group (MNI = 48, 8, 6, $t = 4.14$, $P_{FWE} = 0.013$ SVC, voxels = 7; Figure 3a). We also analyzed the interaction between face types and treatment and the main effect of emotional face processing (all emotional > neutral faces) and no significant effects were observed ($P_{FWE} < 0.05$). Furthermore, examining group differences (OT > PLC) in responses to each face type revealed increased right AI activity in the OT group when presented with neutral (MNI = 48, 8, 6, $t = 4.44$, $P_{FWE} = 0.005$ SVC, voxels = 7) and disgusted faces (MNI = 46, 8, 6, $t = 4.01$, $P_{FWE} = 0.018$ SVC, voxels = 12).

We additionally investigated face emotion-specific effects of OT during the interoceptive task by separately examining contrasts between each emotional face and the neutral one. A greater right amygdala response (MNI = 26, -8, -16, $t = 3.68$, $P_{FWE} = 0.006$ SVC, voxels = 7; Figure 4a) in the OT group only occurred when subjects were presented with disgust faces compared to neutral faces (OT $_{\text{disgust faces}} > \text{neutral faces}$ > PLC $_{\text{disgust faces}} > \text{neutral faces}$). There were no significant effects for other face emotions in the a priori ROIs ($P_{FWE} < 0.05$ SVC). Additional exploratory whole-brain analysis ($P_{FWE} < 0.05$) revealed no significant effects of OT in other regions.

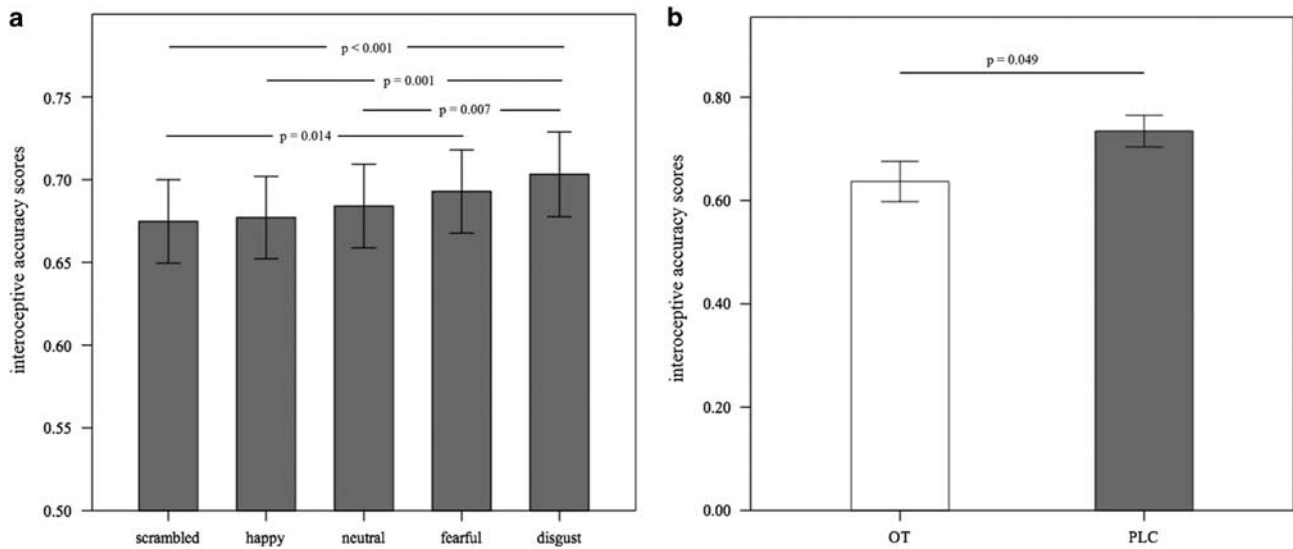


Figure 2 (a) Interoceptive accuracy during presentation of the different face stimuli in Experiment 2 for oxytocin (OT) and placebo (PLC) groups combined. (b) Interoceptive accuracy for OT and PLC groups separately across the different face stimuli in Experiment 2.

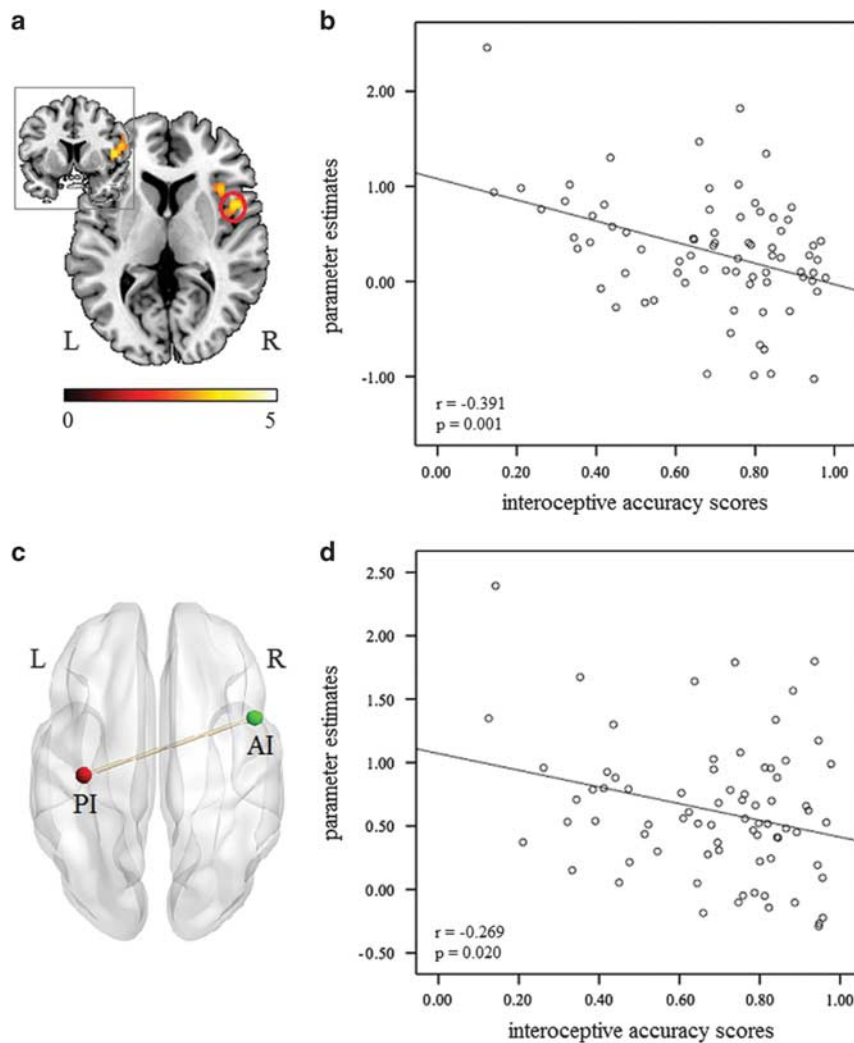


Figure 3 (a) The main effect of treatment (oxytocin > placebo) on the right anterior insula (AI) in Experiment 2. Statistical maps are displayed with a height threshold of $p < 0.005$. (b) Scatter plot showing that right AI activity was correlated with interoceptive accuracy. (c) Oxytocin increased the functional connectivity of the right AI with the left posterior insula (PI). (d) Scatter plot showing that right AI-left PI functional connectivity was correlated with interoceptive accuracy.

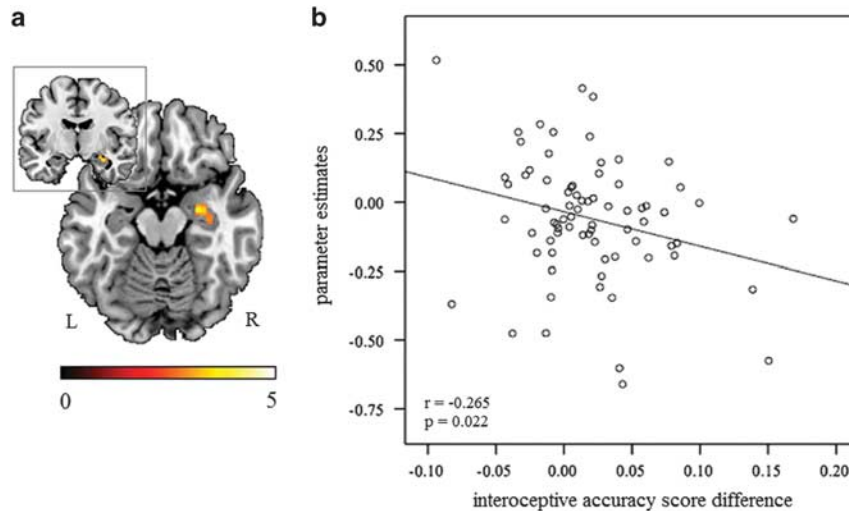


Figure 4 (a) Oxytocin increased the right amygdala activity in response to disgust vs neutral faces (Oxytocin_{disgust faces > neutral faces} > Placebo_{disgust faces > neutral faces}) in Experiment 2. Statistical maps are displayed with a height threshold of $p < 0.005$. (b) Scatter plot showing that right amygdala activity was correlated with the interoceptive accuracy difference score (disgust minus neutral).

Subsequently a PPI analysis was employed to examine whether OT's effects on AI activity were accompanied by effects on its coupling with other regions. In line with behavioral and fMRI findings this analysis focused on treatment main effects. This revealed increased functional connectivity of the right AI (MNI = 48, 8, 6) specifically with the left PI (MNI = -36, -20, 20, $t = 3.71$, $P < 0.001$ uncorrected, voxels = 11; Figure 3c) in the OT compared with the PLC group across different face types. To further confirm that the increased AI-PI connectivity did not represent unspecific effects of OT, but rather develop in interaction with the simultaneously presented social stimuli, the same coordinate was used to run a concordant PPI analysis in Experiment 1. Findings revealed no evidence of treatment effects on functional AI-PI coupling during the heartbeat detection task alone in Experiment 1 ($P < 0.001$ uncorrected).

Brain behavior associations. A correlation analysis between parameter estimates and interoceptive accuracy scores was performed across different face stimuli to further explore associations between OT-induced modulation of neuronal responses and interoceptive accuracy (Supplementary information). There was a significant negative correlation between parameter estimates extracted from the right AI (MNI = 48, 8, 6) and interoceptive accuracy scores across all face types in both the OT (Pearson $r = -0.347$, $df = 37$, $p = 0.036$) and PLC groups (Pearson $r = -0.334$, $df = 38$, $p = 0.040$) alone and combined (Pearson $r = -0.391$, $df = 75$, $p = 0.001$; Figure 3b). There was a similar negative correlation between the AI-PI functional connectivity strength and interoceptive accuracy scores across groups (Pearson $r = -0.269$, $df = 75$, $p = 0.020$; Figure 3d), indicating that greater AI activation and stronger functional connectivity with the PI were both associated with lower interoceptive accuracy. These associations between right AI activity and functional connectivity and interoceptive accuracy were not significant in Experiment 1 ($ps > 0.231$).

For the disgust emotion-specific effect (disgust > neutral), the parameter estimates in the right amygdala (MNI = 26,

-8, -16) were also significantly correlated with the interoceptive accuracy difference scores across groups (Pearson $r = -0.265$, $df = 75$, $p = 0.022$; Figure 4b), indicating that increased amygdala, emotion-specific responses were associated with decreased interoceptive accuracy.

DISCUSSION

The present study first investigated whether intranasal OT influences behavioral responses and neural circuitry involved in interoceptive awareness, using a heartbeat detection task, and second whether this would be affected by simultaneous exposure to salient social cues. Our results do not provide any evidence that OT significantly influences interoceptive awareness (in terms of interoceptive accuracy) *per se* or associated neural responses in the AI and dACC. On the other hand, when subjects were simultaneously presented with face stimuli, OT both decreased interoceptive accuracy and increased right AI activation and its functional connectivity with the PI. These findings are consistent with OT promoting a switch of attention away from interoceptive cues towards external social salient ones.

In agreement with previous studies (Critchley *et al*, 2004; Pollatos *et al*, 2007b; Kurth *et al*, 2010; Chang *et al*, 2012; Zaki *et al*, 2012; Uddin *et al*, 2014), in Experiment 1 we found increased activity in bilateral AI, PI, and dACC in a heartbeat detection task. However, contrary to our original hypothesis intranasal OT neither altered interoceptive accuracy nor associated neural activity in the interoceptive networks. Thus, we found no evidence to support the view that OT can influence interoceptive processing *per se*.

When subjects were presented with neutral and emotional face stimuli in Experiment 2, interoceptive accuracy was higher when they viewed negative emotional faces, especially disgust faces, than with neutral, scrambled or happy ones. Thus, negative emotional stimuli may produce a greater facilitation of interoceptive accuracy, which coincides with previous findings that interoceptive accuracy is closely associated with the intensity of negative but not positive emotion (Critchley *et al*, 2004) and that patients

characterized by exaggerated negative affective states exhibit higher interoceptive sensitivity (Ehlers and Breuer, 1996; Dunn *et al*, 2010). However, OT treatment produced a main effect in reducing interoceptive accuracy across different face stimuli. Furthermore, this effect of OT was associated both with increased right AI activity and with increased functional connectivity between the right AI and the left PI. Importantly in this context, both right AI activity and its functional connectivity with the PI were negatively correlated with interoceptive accuracy across groups whereas these associations were absent in Experiment 1 when subjects were not presented with face stimuli. Thus the simultaneous presentation of external stimuli interfered with the role of AI in processing interoceptive signals, with the correlation indicating that a stronger engagement of AI in processing external salient cues led to lower interoceptive accuracy.

Several previous studies have reported considerable overlap between AI activation following interoception- and emotion-based tasks (Critchley *et al*, 2004; Pollatos *et al*, 2007b; Zaki *et al*, 2012). As Experiment 1 showed no evidence that OT enhanced right AI activity during interoceptive processing alone, the fact that it did do so in Experiment 2 indicates that this was primarily a result of facilitation of the salience of all the face stimuli presented. While we had no behavioral measure of such an OT-enhancement of the salience of these face stimuli, numerous previous studies have demonstrated evidence that OT increases attention towards (Domes *et al*, 2013; Xu *et al*, 2015), or recognition of (Fischer-Shofty *et al*, 2010; Shahrestani *et al*, 2013), neutral, positive or negative faces using different paradigms. Indeed, one of the most dominant current theories of OT's function is that it enhances social salience (Shamay-Tsoory and Abu-Akel, 2016), and following intranasal OT the AI has consistently been found to have altered activity and/or functional connectivity with other regions of the salience network during various social tasks (Riem *et al*, 2011; Striepens *et al*, 2012; Gao *et al*, 2016). Thus, OT may facilitate a shift of information integration in the AI away from interoceptive feedback-processing towards external social stimuli in order to guide behavior during social interactions, especially when confronted with highly salient social cues such as faces. This also resonates more generally with the role of the AI as a core hub involved in switching between external task-oriented and self-oriented networks (Menon and Uddin, 2010; Uddin, 2015). It is notable in this respect that OT only influenced the right AI which exhibits the highest level of functional diversity and connectivity with other cortical regions and is of most importance in terms of salience processing (Uddin, 2015).

Interestingly, OT specifically increased functional connectivity between the right AI and the left PI during presentation of face stimuli in Experiment 2. The PI is proposed to process primary sensory components of interoceptive signals (Chang *et al*, 2012; Uddin *et al*, 2014), whereas the AI integrates information from the PI and generates interoceptive awareness (Craig, 2002). Thus, OT may facilitate regulatory control of the sensory interoceptive information from the PI by the right AI as a result of its enhanced engagement in processing external salient cues. This idea is supported by the observation that the AI-PI functional connectivity was only correlated with interoceptive accuracy in the presence of external social stimuli.

However, given the exploratory nature and liberal threshold used in this analysis, inferences regarding this finding should be drawn with caution.

Overall, our findings coincide with a role of OT in decreasing self-orientated behavior (internal) and increasing other-oriented (external) responses (Abu-Akel *et al*, 2015; Bartz *et al*, 2015; Zhao *et al*, 2016). They also provide support for the hypothesis of OT function in interoceptive inference and, in particular, failures of interoceptive inference in conditions like autism and schizophrenia (Klein *et al*, 2013; Quattrocki and Friston, 2014), by showing experimentally that an OT-enhancement of attention to socially salient stimuli via the right AI, decreases corresponding sensitivity to interoceptive information. In predictive coding models, attention to interoceptive or exteroceptive modalities is mediated by the precision of ascending prediction errors encoded by the gain of postsynaptic responses in cells signaling prediction errors (such as superficial pyramidal cells and von Economo cells in the insular cortex). Increasing the precision of prediction errors high in the hierarchy (eg, in the AI) is equivalent to decreasing the precision at lower, sensory, levels (such as the PI). Thus, OT may increase the gain of neuronal populations higher in the hierarchy, resulting in a failure to attend to interoceptive signals and decreased interoceptive accuracy. Failure of attention or sensory attenuation in interoceptive inference may contribute to autism and schizophrenia through false inferences (eg, illusions, delusions, and hallucinations) about bodily states and, indeed, other modalities. Autistic individuals have problems with detecting and responding appropriately to external social cues but can sustain better attention to interoceptive cues over long periods of time similar to those the present study used (Schauder *et al*, 2015). However, improved interoceptive accuracy is negatively associated with performance on the rubber hand illusion task, a measure of external body awareness, indicating that the increased focus on interoceptive cues resulted in a decreased focus on exteroceptive ones (Schauder *et al*, 2015). It is notable in this respect that right AI activation is significantly decreased in autistic individuals when they try to infer social emotions in others, and that this is remedied by intranasal OT treatment (Aoki *et al*, 2014). The AI-PI functional connectivity is also weaker in ASD individuals (Ebisch *et al*, 2011), which might possibly indicate a reduced regulatory control of incoming interoceptive information when external social salient cues are present.

Individuals with emotional disorders exhibit higher interoceptive sensitivity (Ehlers and Breuer, 1996; Dunn *et al*, 2010), which might even represent an etiological factor for the conditions (Paulus and Stein, 2010; Garfinkel and Critchley, 2013). The finding that OT decreased interoceptive accuracy may therefore also indicate that it can exert an anxiolytic effect in such disorders by reducing sensitivity to interoceptive signals.

In line with our main hypothesis, we primarily found main effects of treatment across different facial stimuli, although OT did particularly increase activity in the right AI in response to both neutral and disgust faces. A previous study has reported that OT can increase attention to neutral faces (Xu *et al*, 2015) and effects of disgust faces may reflect the key role of AI in processing this emotion (Phillips *et al*, 1997). In addition, we did find some evidence for emotion-

specific effects of OT in the amygdala, with enhanced responses to disgust vs neutral faces. The amygdala is both implicated in disgust emotion processing (Fitzgerald *et al*, 2006) and OT's effects on processing of social cues (Striepens *et al*, 2011; Gao *et al*, 2016). These amygdala responses were negatively correlated with interoceptive accuracy across groups, indicating that enhanced emotional encoding in the amygdala led to reduced interoceptive awareness. This provides further support for OT acting primarily on salience processing to reduce attention to interoceptive cues.

The current study has several limitations which should be acknowledged. First, as we only used face stimuli in Experiment 2 we do not know if OT's effects on interoceptive accuracy and neural networks might also extend to non-social stimuli. Second, we only used male subjects in order to avoid having to consider possible menstrual cycle effects, but there have been numerous previous studies reporting sex-dependent effects of OT (Gao *et al*, 2016).

In conclusion, the present study provides both behavioral and neural evidence that OT switches attention away from interoceptive cues towards external social salient ones. Thus OT may promote social interactions by increasing attention towards awareness of others more than self by modulating right AI circuitry involved in both interoceptive and salience processing. These findings suggest that OT may have therapeutic potential in emotional disorders where the interplay between interoceptive and exteroceptive processing of emotional stimuli is impaired.

FUNDING AND DISCLOSURE

This study was supported by the National Natural Science Foundation of China (NSFC) grant (grant number 31530032; 91632117) and the Fundamental Research Funds for the Central Universities (ZYGX2015Z002).

REFERENCES

- Abu-Akel A, Palgi S, Klein E, Decety J, Shamay-Tsoory S (2015). Oxytocin increases empathy to pain when adopting the other-but not the self-perspective. *Soc Neurosci* **10**: 7–15.
- Aoki Y, Yahata N, Watanabe T, Takano Y, Kawakubo Y, Kuwabara H *et al* (2014). Oxytocin improves behavioural and neural deficits in inferring others' social emotions in autism. *Brain* **137**: 3073–3086.
- Bartz JA, Lydon JE, Kolevzon A, Zaki J, Hollander E, Ludwig N *et al* (2015). Differential effects of oxytocin on agency and communion for anxiously and avoidantly attached individuals. *Psychol Sci* **26**: 1177–11860.
- Bartz JA, Zaki J, Bolger N, Ochsner KN (2011). Social effects of oxytocin in humans: context and person matter. *Trends Cogn Sci* **15**: 301–309.
- Brewer R, Happé F, Cook R, Bird G (2015). Commentary on "Autism, oxytocin and interoception": alixithymia, not autism spectrum disorders, is a consequence of interoceptive failure. *Neurosci Biobehav Rev* **56**: 348–353.
- Chang LJ, Yarkoni T, Khaw MW, Sanfey AG (2012). Decoding the role of the insula in human cognition: functional parcellation and large-scale reverse inference. *Cereb Cortex* **23**: 739–749.
- Craig AD (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nat Rev Neurosci* **3**: 655–666.
- Critchley HD, Harrison NA (2013). Visceral influences on brain and behavior. *Neuron* **77**: 624–638.
- Critchley HD, Wiens S, Rotshtein P, Öhman A, Dolan RJ (2004). Neural systems supporting interoceptive awareness. *Nat Neurosci* **7**: 189–195.
- Damasio AR (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. Grosset/Putnam Books: New York.
- Domes G, Sibold M, Schulze L, Lischke A, Herpertz SC, Heinrichs M (2013). Intranasal oxytocin increases covert attention to positive social cues. *Psychol Med* **43**: 1747–1753.
- Dunn BD, Stefanovitch I, Evans D, Oliver C, Hawkins A, Dalgleish T (2010). Can you feel the beat? Interoceptive awareness is an interactive function of anxiety- and depression-specific symptom dimensions. *Behav Res Ther* **48**: 1133–1138.
- Ebisch SJH, Gallese V, Willems RM, Mantini D, Groen WB, Romani GL *et al* (2011). Altered intrinsic functional connectivity of anterior and posterior insula regions in high-functioning participants with Autism Spectrum Disorder. *Hum Brain Mapp* **32**: 1013–1028.
- Ehlers A, Breuer P (1996). How good are patients with panic disorder at perceiving their heartbeats? *Biol Psychol* **42**: 165–182.
- Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K *et al* (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage* **25**: 1325–1335.
- Fischer-Shofty M, Shamay-Tsoory SG, Harari H, Levkovitz Y (2010). The effect of intranasal administration of oxytocin on fear recognition. *Neuropsychologia* **48**: 179–184.
- Fitzgerald DA, Angstadt M, Jelsone LM, Nathan PJ, Phan KL (2006). Beyond threat: amygdala reactivity across multiple expressions of facial affect. *NeuroImage* **30**: 1441–1448.
- Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RS (1994). Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp* **2**: 189–210.
- Gao S, Becker B, Luo L, Geng Y, Zhao W, Yin Y *et al* (2016). Oxytocin, the peptide that bonds the sexes also divides them. *Proc Natl Acad Sci USA* **113**: 7650–7654.
- Garfinkel SN, Critchley HD (2013). Interoception, emotion and brain: new insights link internal physiology to social behaviour. Commentary on: "anterior insular cortex mediates bodily sensibility and social anxiety" by Terasawa *et al*. (2012). *Soc Cogn Affect Neurosci* **8**: 231–234.
- Gimpl G, Fahrenholz F (2001). The oxytocin receptor system: structure, function, and regulation. *Physiol Rev* **81**: 629–683.
- Guastella AJ, Hickie IB, McGuinness MM, Otis M, Woods EA, Disinger HM *et al* (2013). Recommendations for the standardisation of oxytocin nasal administration and guidelines for its reporting in human research. *Psychoneuroendocrinology* **38**: 612–625.
- Guastella AJ, Mitchell PB, Matthews F (2008). Oxytocin enhances the encoding of positive social memories in humans. *Biol Psychiatry* **64**: 256–258.
- Hurlemann R, Patin A, Onur OA, Cohen MX, Baumgartner T, Metzler S *et al* (2010). Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. *J Neurosci* **30**: 4999–5007.
- Jankowski M, Danalache B, Wang D, Bhat P, Hajjar F, Marcinkiewicz M *et al* (2004). Oxytocin in cardiac ontogeny. *Proc Natl Acad Sci USA* **101**: 13074–13079.
- Kemp AH, Quintana DS, Kuhnert RL, Griffiths K, Hickie IB, Guastella AJ (2012). Oxytocin increases heart rate variability in humans at rest: implications for social-approach-related motivation and capacity for social engagement. *PLoS ONE* **7**: e44014.
- Klein TA, Ullsperger M, Danielmeier C (2013). Error awareness and the insula: links to neurological and psychiatric diseases. *Front Hum Neurosci* **7**: 14.
- Kurth F, Zilles K, Fox PT, Laird AR, Eickhoff SB (2010). A link between the systems: functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Struct Funct* **214**: 519–534.

- Lange CG, James W (1967). *The Emotions (Reprinted)*. In: Dunlap Keditor Hafner Publishing Co: New York.
- McLaren DG, Ries ML, Xu G, Johnson SC (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): a comparison to standard approaches. *NeuroImage* **61**: 1277–1286.
- Menon V, Uddin LQ (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Struct Funct* **214**: 655–667.
- Paulus MP, Stein MB (2010). Interoception in anxiety and depression. *Brain Struct Funct* **214**: 451–463.
- Phillips ML, Young AW, Senior C, Brammer M, Andrew C, Calder AJ et al (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature* **389**: 495–498.
- Pollatos O, Gramann K, Schandry R (2007a). Neural systems connecting interoceptive awareness and feelings. *Hum Brain Mapp* **28**: 9–18.
- Pollatos O, Kirsch W, Schandry R (2005). On the relationship between interoceptive awareness, emotional experience, and brain processes. *Cogn Brain Res* **25**: 948–962.
- Pollatos O, Schandry R, Auer DP, Kaufmann C (2007b). Brain structures mediating cardiovascular arousal and interoceptive awareness. *Brain Res* **1141**: 178–187.
- Quattrocki E, Friston K (2014). Autism, oxytocin and interoception. *Neurosci Biobehav Rev* **47**: 410–430.
- Riem MM, Bakermans-Kranenburg MJ, Pieper S, Tops M, Boksem MA, Vermeiren RR et al (2011). Oxytocin modulates amygdala, insula, and inferior frontal gyrus responses to infant crying: a randomized controlled trial. *Biol Psychiatry* **70**: 291–297.
- Schauder KB, Mash LE, Bryant LK, Cascio CJ (2015). Interoceptive ability and body awareness in autism spectrum disorder. *J Exp Child Psychol* **131**: 193–200.
- Seth AK (2013). Interoceptive inference, emotion, and the embodied self. *Trends Cogn Sci* **17**: 565–573.
- Shahrestani S., Kemp AH, Guastella AJ (2013). The impact of a single administration of intranasal oxytocin on the recognition of basic emotions in humans: a meta-analysis. *Neuropsychopharmacology* **38**: 1929–1936.
- Shamay-Tsoory SG, Abu-Akel A (2016). The social salience hypothesis of oxytocin. *Biol Psychiatry* **79**: 194–202.
- Striepens N, Kendrick KM, Maier W, Hurlmann R (2011). Prosocial effects of oxytocin and clinical evidence for its therapeutic potential. *Front Neuroendocrinol* **32**: 426–450.
- Striepens N, Scheele D, Kendrick KM, Becker B, Schäfer L, Schwalba K et al (2012). Oxytocin facilitates protective responses to aversive social stimuli in males. *Proc Natl Acad Sci USA* **109**: 18144–18149.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N et al (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* **15**: 273–289.
- Uddin LQ (2015). Saliency processing and insular cortical function and dysfunction. *Nat Rev Neurosci* **16**: 55–61.
- Uddin LQ, Kinnison J, Pessoa L, Anderson ML (2014). Beyond the tripartite cognition-emotion-interoception model of the human insular cortex. *J Cogn Neurosci* **26**: 16–27.
- Xu L, Ma X, Zhao W, Luo L, Yao S, Kendrick KM (2015). Oxytocin enhances attentional bias for neutral and positive expression faces in individuals with higher autistic traits. *Psychoneuroendocrinology* **62**: 352–358.
- Zaki J, Davis JI, Ochsner KN (2012). Overlapping activity in anterior insula during interoception and emotional experience. *NeuroImage* **62**: 493–499.
- Zhao W, Yao S, Li Q, Geng Y, Ma X, Luo L et al (2016). Oxytocin blurs the self-other distinction during trait judgments and reduces medial prefrontal cortex responses. *Hum Brain Mapp* **37**: 2512–2527.

Supplementary Information accompanies the paper on the Neuropsychopharmacology website (<http://www.nature.com/npp>)