OXFORD

Cerebral Cortex, 2022, 1–18
https://doi.org/10.1093/cercor/bhac167
Original Article

The role of oxytocin in modulating self–other distinction in human brain: a pharmacological fMRI study

Yuanchen Wang, BA, BS[1,2,3], Ruien Wang[1,2], Haiyan Wu, PhD[1,2,*]

[1]Centre for Cognitive and Brain Sciences, N21 Research Building, University of Macau, Avenida da Universidade, Taipa, Macau SAR 999078 , China,
[2]Department of Psychology, E21B Humanities and Social Sciences Building, University of Macau, Avenida da Universidade, Taipa, Macau SAR 999078, China,
[3]Department of Biomedical Informatics, Harvard Medical School, 10 Shattuck Street, Boston, MA 02115, USA

*Corresponding author: Centre for Cognitive and Brain Sciences and Department of Psychology, University of Macau, N21 Research Building, Avenida da Universidade, Taipa, Macau SAR 999078, China. Email: haiyanwu@um.edu.mo

Self–other distinction is crucial for human interaction. Although with conflicting results, studies have found that oxytocin (OT) sharpens the self–other perceptual boundary. However, little is known about the effect of OT on self–other perception, especially its neural basis. Moreover, it is unclear whether OT influences self–other discrimination when the other is a child or an adult. This double-blind, placebo-controlled study investigated the effect of OT on self-face perception at the behavioral and neural levels. For the stimuli, we morphed participants' faces and child or adult strangers' faces, resulting in 4 conditions. After treatment with either OT or placebo, participants reported whether a stimulus resembled themselves while being scanned using functional magnetic resonance imaging (fMRI). Behavioral results showed that people judged adult-morphed faces better than child-morphed faces. Moreover, fMRI results showed that the OT group exhibited increased activity in visual areas and the inferior frontal gyrus for self-faces. This difference was more pronounced in the adult-face condition. In multivariate fMRI and region of interest analyses, better performance in the OT group indicated that OT increased self–other distinction, especially for adult faces and in the left hemisphere. Our study shows a significant effect of OT on self-referential processes, proving the potential effect of OT on a left hemisphere self-network.

Key words: face; fMRI; oxytocin; self–other distinction; self-resemblance.

Introduction

Self-related and other-related information processing, such as the self–other distinction and in-group–out-group separation, greatly contributes to everyday social decisions and interactions. When making self-related or other-related judgments, one of the most important cues is facial stimuli. Humans infer genetic relatedness through the resemblance between presented faces and self-face to adjust their altruism or investment in others' behaviors (Alvergne et al. 2007; Wu et al. 2013). Although some studies have shown individual differences in processing self-resemblance faces (Brédarta and French 1999; Bressan and Dal Martello 2002; Alvergne et al. 2007; Wu et al. 2013), these faces also increased participants' trust toward individuals by evoking the feeling of being close (DeBruine 2002). Moreover, Platek and Kemp (2009) also confirmed that implicit trust evaluation of self-resemblance faces would activate the reward-related brain areas, suggesting an important role of face perception in social interaction.

Successfully identifying and distinguishing between information related to oneself and others play a fundamental role in social life (Lamm et al. 2016). Except for the variance in stimuli (faces) themselves, the administration of neuropeptides could also affect people's related performance. For example, studies have shown that oxytocin (OT) promotes social connections and improves social interactions, such as trust and empathy (Baumgartner et al. 2008; Hurlemann et al. 2010). However, it is unclear how OT modulates these mechanisms and impacts the neural representation of social relationships. One possibility is that OT decreases the self–other boundary and, in turn, increases social interactions. To further understand the mechanism of OT's effect of OT on social behaviors, it is crucial to investigate the effect of OT based on both behavioral and neural responses during self–other face distinction tasks.

Given that OT can modulate the salient detection of social stimuli (Shamay-Tsoory and Abu-Akel 2016; Tillman et al. 2019), it is expected that OT influences face processing at both the behavioral and neural levels (Bate et al. 2014, 2015). For example, research has shown that OT enhances recognition memory for faces but not for nonfacial stimuli (Rimmele et al. 2009). Nevertheless, more recent evidence suggests that this enhancement is attributable to participants' tendency to classify unfamiliar faces as familiar faces rather than to an improvement in recognition memory for faces, as indexed by heightened sensitivity based on signal detection theory (SDT) (Bate et al. 2015). In addition, OT decreases amygdala responses to emotional faces, suggesting that

Received: February 1, 2022. **Revised:** April 3, 2022. **Accepted:** April 8, 2022

OT tends to exert a greater impact on more socially salient information (Domes et al. 2007). Altogether, these findings corroborate the role played by OT in increasing the social salience of facial features in face processing, which may in turn influence the representation of social relationships and further social behaviors.

Previous studies have shown that affiliative behaviors are supported by neural mechanisms associated with the social brain, including the medial prefrontal cortex (MPFC), anterior cingulate cortex (ACC), temporal parietal junction (TPJ), inferior frontal gyrus (IFG), and anterior insula (Frith 2007; Adolphs 2009; Barrett and Satpute 2013; Stanley and Adolphs 2013; Martin et al. 2019). These brain areas are activated when individuals construct representations of relationships between themselves and others and use this information to understand and guide social behavior (Adolphs 2001, 2010; Gobbini et al. 2011). Therefore, it is possible that interpersonal psychological distance and the self–other distinction are mediated by these social brain regions. Furthermore, the dorsomedial prefrontal cortex (dmPFC), a core region of the social brain, is involved in mentalization and social cognition (Denny et al. 2012; Wu, Liu, et al. 2020b). Additionally, activity in this brain region has been linked to the self–other distinction (Amodio and Frith 2006), and it is thought to depend on how close we perceive other individuals and how similar they feel to them (Krienen et al. 2010). Specifically, researchers have shown that the dmPFC is activated when we make inferences about the mental state of dissimilar others compared to similar others (Mitchell et al. 2006). Thus, the dmPFC appears to be a core region that mediates social relationships and represents the mental state of other individuals, particularly when they are dissimilar or unfamiliar. Similar to the dmPFC, the ACC, posterior cingulate cortex, and TPJ have also been shown to play a role in mental state reasoning (Adolphs 2001, 2003; Wu, Liu, et al. 2020b). Previous studies have shown that these regions can represent the position of the body in space and help determine where an individual looks at (Frith and Frith 2006). ACC was also found to be active in self-monitoring behaviors, such as recognizing oneself and others (Gallagher and Frith 2003; Frith and Frith 2006). Together, these brain regions may serve as neural underpinnings of the self–other distinction.

As previously mentioned, one's own face is considered a salient self-related stimulus; therefore, it has long been applied in self-recognition and self–other perceptual difference studies (Keenan et al. 2000; Tsakiris 2008). Researchers have shown that OT can modulate self-resemblance face processing (Colonnello et al. 2013; Zhao et al. 2016) but with conflicting results. For example, with self-stranger face morphing, a behavioral study indicated that OT increased the ability to recognize differences between self and others in the self–other face differentiation task and increased positive evaluation of others (Colonnello et al. 2013). However, another study indicated that OT blurs the self–other distinction and reduces the mPFC activity during self-trait judgments (Zhao et al. 2016). Another important factor that influences the effect of OT is the age of the self-resembling faces. Based on a functional magnetic resonance imaging (fMRI) study, child and adult faces were found to have different activation levels despite similar activation regions (Marusak et al. 2013). Therefore, it is also crucial to consider the age factor in the current experiment and the faces of both adults and children involved.

Based on the current research findings on the effect of OT on self–other distinction and the age factor, 2 different hypotheses can be generated. First, the effect of OT on the self–other distinction may be similar for both adult and child faces. Self-resembling faces indicate genetic relatedness and higher trustworthiness, which would activate reward-related brain regions such as the ventral superior frontal gyrus, right ventral IFG, and left medial frontal gyrus (MFG) (Platek and Kemp 2009). However, other evidence has shown that OT increases self–other differentiation on peer-age faces (i.e. adult faces) but may not be evident for child faces. For instance, our previous study indicated that males are more sensitive to self-morphed adult faces than to self-morphed child faces (Wu et al. 2013). Thus, it is also possible that OT increases self–other discrimination more under adult-face conditions than under child-face conditions.

The current study used a pharmacological fMRI approach to examine the neural correlates of the effects of OT on the self–other distinction using self-morphed adult and child faces. In this study, we collected psychometric data on participants' personality traits, behavioral data for the self–other discrimination task (e.g. accuracy and reaction time [RT]), and fMRI data of the participants while executing the task. According to the social salience hypothesis of OT and own-age bias in face perception, we hypothesized that OT would affect the self–other distinction and would be different for adult and child faces. In addition, we expected that OT would modulate face processing and self-processing brain activity depending on self-resemblance. Finally, we hypothesized that the effect of OT on the self–other differentiation task (self vs. other) may be associated with one's different personality traits.

Methods

Participants

We recruited 59 healthy, right-handed, male participants (age: 20.9 ± 2.32 years) with 13–18 years of education. They participated in this study through an online recruitment system. All participants filled out a screening form and were included in the study only if they confirmed that they were not suffering from any significant medical or psychiatric illness, were not currently using medication, and were not consuming alcohol nor smoking on a daily basis. The participants were instructed to refrain from smoking or drinking (except for water) for 2 h before the experiment. The participants received a full

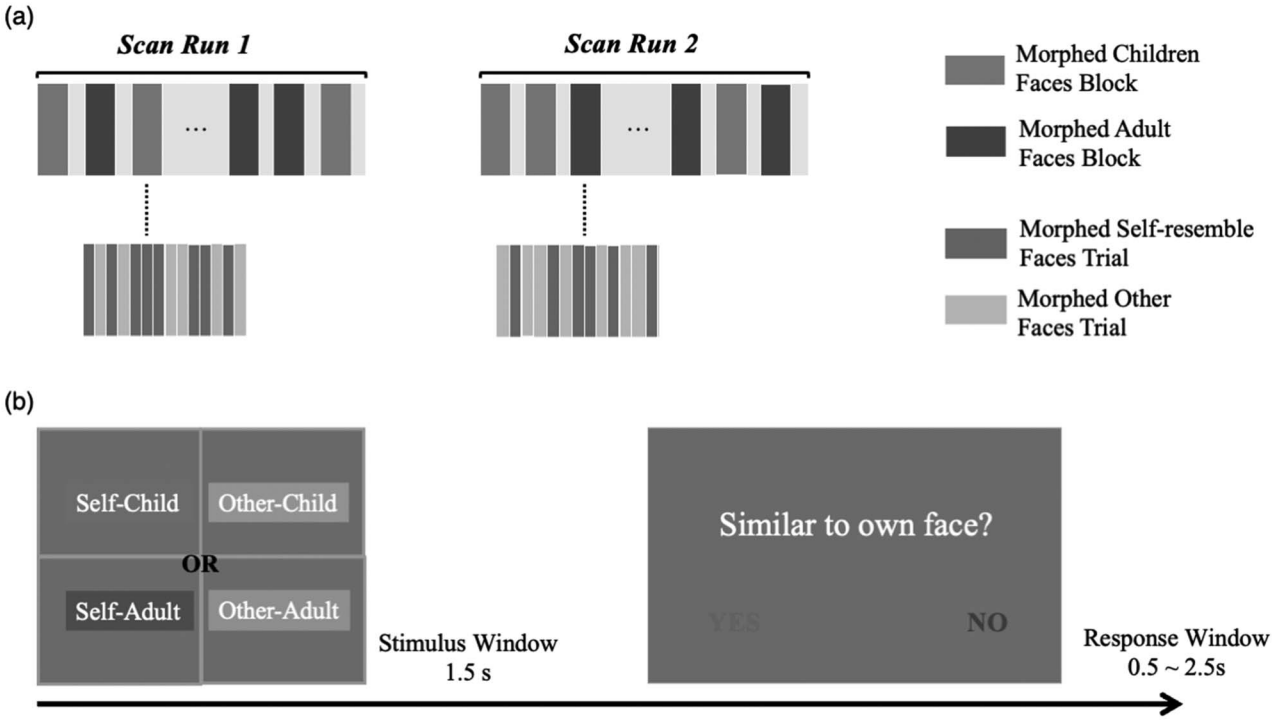

Fig. 1. Sequence and conditions of the experiment. a) Time sequence for the whole experiment. b) Time sequence for a single stimuli in each block. Different conditions are color coded indicated by legend in the upper right corner. Faces shown in the self-child condition were morphed using the participant's own face and a stranger child's face; faces in the other-child condition were morphed using an adult stranger's face (the same gender as the participant) and a stranger child's face; faces in the self-adult condition were morphed using the participant's own face and a stranger adult's face (the same gender as the participant); faces in the other-adult condition were morphed using 2 adult strangers' faces (the same gender as the participant).

debriefing on the completion of the experiment. Written informed consent was obtained from each participant before the experiment. This study was approved by the local ethics committee.

Task design

In this study, we divided the participants into 2 groups: the OT group, which was treated with nasal OT, and the placebo (PL) group, which was given the same volume of saline water via nasal administration as the placebo. For the stimuli presented to each participant, we created 4 experimental conditions (self-child, self-adult, other-child, and other-adult) by morphing the participant's face with 1 of 2 adult faces with neutral expression and a 1.5-year-old child face. Figure 1 shows an illustration of the morphed faces and the resulting 4 experimental conditions, a process similar to that of our previous study (Wu et al. 2013).

We collected data in two 6-min runs for each participant. In each run, there were 6 blocks, including 3 adult and 3 child blocks, which were presented randomly. In the adult blocks, the stimuli shown were those morphed with adults: the self-adult and other-adult conditions. Similarly, only the self-child and other-child conditions were included in the child blocks. For each trial in the study, the facial stimuli were presented for 1,500 ms, which showed the target morphed face. The participants were asked to judge whether the face resembled their face in the following response window (Fig. 1).

Material preparation and acquisition

A full-face photograph of each participant was taken 1 week before the formal study. Photographs were also taken 3 days before the scanning day. The participants were asked to maintain a neutral expression when facing the camera. Morphed faces were created based on the photographs taken. To exclude the gender effect of the faces, all faces used for morphing were of the same gender as that of the participants. All faces were processed with Adobe Photoshop CS to standardize the picture to black and white, with only the interior characteristics of the face being retained. Abrosoft Fanta Morph (www.fantamorph.com) software was then used to create the 50/50 morph of the 2 selected faces, a method similar to that used in previous studies (Platek et al. 2004, 2005; Platek and Kemp 2009). Thirty calibration locations were used to create the morphed face in a standard face space, and all output morphed faces were resized to 300 × 300 dpi. All stimuli were presented on a 17-inch Dell monitor with a screen resolution of 1,024 × 768 pixels and a 60 Hz refresh frequency. The visual angle of the face images was 4.3° × 4.6°, and the mean luminance of the stimulus was 166 cd/m^2.

Psychological scales

We collected demographic and psychometric data of all participants. The measures were predominantly recorded before the experiment, with several exceptions.

Table 1. Psychometric data and questionnaire scores.

	PL	OT	*t*-score	*P*-value
PA	31.76 (5.33)	33.1 (6.99)	−0.8242	0.4136
NA	18.17 (6.12)	17.17 (6.02)	0.6272	0.5331
SA	37.24 (9.29)	35.03 (8.88)	0.9248	0.3591
PA after	29.38 (5.73)	31.21 (6.86)	−1.1009	0.2758
NA after	17.48 (6.65)	16.69 (6.2)	0.4697	0.6404
SA after	38.69 (9.1)	37.34 (10.73)	0.5148	0.6088
AQ	22.14 (4.84)	23.59 (5.64)	−1.0494	0.2986
Neuroticism	44.97 (12.01)	40.31 (11.75)	1.492	0.1413
Extraversion	48.9 (10.3)	51.31 (10.42)	−0.8873	0.3787
Openness	55.86 (8.48)	54.69 (7.97)	0.5426	0.5896
Agreeableness	61.86 (6.58)	59.24 (8)	1.3619	0.1789
Conscientiousness	60.52 (11.66)	63.52 (10.18)	−1.0439	0.3011
IRI	96.76 (9.64)	92.17 (10.17)	1.7619	0.0836
TAS-20	53.55 (9.3)	49.48 (9.54)	1.6443	0.1057
BIS all	66.59 (8.83)	65.59 (9.81)	0.4079	0.6849
SES score	30.03 (4.36)	31.62 (3.86)	−1.4666	0.1482
NPI score	121.24 (25.38)	127.07 (25.25)	−0.8766	0.3845
SDO score	53.45 (12.45)	55.72 (10.53)	−0.7517	0.4555
FNE short	42.28 (7.43)	39.93 (7.45)	1.2001	0.2351
FNE all items	103.86 (15.44)	97.52 (15.84)	1.545	0.128
Mav score	−3.34 (11.53)	1.76 (14.14)	−1.5064	0.1378
Father love	23.34 (9.42)	22.66 (7.58)	0.3072	0.7598
Mother love	23.03 (9.58)	24.93 (7.64)	−0.8334	0.4083
NFCC	163.24 (16.36)	160.93 (18.92)	0.4974	0.6209
IUS All	74 (18.54)	75.55 (16.48)	−0.3369	0.7374
IUS Chinese	66.83 (16.04)	67.69 (14.14)	−0.2171	0.829
PT	22.48 (4.04)	22.34 (2.32)	0.1594	0.8741
FS Scale	25.17 (5.52)	23.14 (5.4)	1.4188	0.1615
Empathic concern	26.76 (3.66)	25.28 (4.84)	1.3158	0.194
PD	22.34 (3.92)	21.41 (4.37)	0.8538	0.3969

Note: Means (standard deviations), and 2-tailed *t*-test (direction: OT − PL) results of the comparisons on demographics data. As none of the comparisons were significant, multiple comparison corrections were not applied. PA, positive affect; NA, negative affect; SA, social anxiety; AQ, autism quotient; IRI, interpersonal reactivity index; TAS-20, Toronto Alexithymia Scale; BIS, behavioral inhibition system scales; SES, socioeconomic status; NPI, Narcissistic Personality Inventory; SDO, social dominance orientation; FNE, fear of negative evaluation; Mav score, Machiavellianism Scale; NFCC, need for closure scale; IUS, intolerance of uncertainty scale. Details are provided in the Supplemental Information.

"PA after," "NA after," and" SA after" are the 3 measures taken after OT/PL administration. These measures used the same scales as PA (positive affect), NA (negative affect), and SA (social anxiety), which were taken before the treatment (Table 1). The psychometric data were collected as a composite of control factors to make sure that people administered OT have no significant difference from people administered PL. For example, interpersonal reactivity index (IRI) is a widely used assessment of empathy (Davis 1983). There are 4 subscales in IRI, including perspective taking (PT), fantasy (FS), empathetic concern (EC), and personal distress (PD). Each subscale includes 7 questions. EC measures individuals' feelings of compassion and concern for others. FS describes the tendency that respondents transpose themselves into fictional characters. PD indicates the extent that individuals feel uneasiness when exposed to the negative experiences of others. PT assesses unplanned attempts to adopt others' points of view. Mean scores were subsequently compared across treatment groups (OT vs. PL) to rule out effects of OT on these measures. Because of our randomized design, we predicted that there would be no significant difference between OT and PL group in psychometric scores.

fMRI image collection

All images were acquired on a 3T Siemens Tim Trio scanner with 12-channel head coil. Functional images employed a gradient-echo echo-planar imaging sequence with following MRI scanner parameters: time echo (TE) = 40 ms, time repetition (TR) = 2 s, flip = 90°, field of view (FOV) = 210 mm, 128 by 128 matrix, 25 contiguous 5 mm slices parallel to the hippocampus, and interleaved. We also acquired the whole-brain T1-weighed anatomical reference images from all participants (TE = 2.15 ms, TR = 1.9 s, flip = 9°, FOV = 256 mm, 176 sagittal slices, slice thickness = 1 mm, perpendicular to the anterior–posterior commissure line).

fMRI imaging analysis

fMRI data preprocessing was performed using Statistical Parametric Mapping software (SPM12: Wellcome Trust Centre for Neuroimaging, London, United Kingdom). The functional image time series were preprocessed to compensate for slice-dependent time shifts, motion corrected, and linearly detrended, then coregistered to the anatomical image, spatial normalized to Montreal Neurological Institute (MNI) space

(http://www.bic.mni.mcgill.ca/ServicesAtlases/HomePage), and spatially smoothed by convolution with an isotropic Gaussian kernel (full-width at half-maximum = 6 mm). The fMRI data were high-pass filtered with a cutoff of 0.01 Hz. The white matter (WM) signal, cerebrospinal fluid signal, and global signal, as well as the 6-dimensional head motion realignment parameters, the realignment parameters squared, their derivatives, and the squared of the derivatives were regressed. The resulting residuals were then low-pass filtered with a cutoff of 0.1 Hz.

Univariate voxel-based analysis and ROI-based analysis

Univariate analysis was conducted using the SPM toolbox with general linear models (GLMs) (Friston et al. 1994). In the GLM analysis, face stimuli blocks were modeled with a boxcar function and convolved with a standard hemodynamic response function. Four conditions were defined by separate regressors: self-adult, other-adult, self-child, and other-child. Six head movement parameters from the spatial realignment were entered as covariates of no interest. Statistical parametric maps were generated for each subject from linear contrasts between each of the 4 conditions. For the ROI analysis, coordinates of cuneus, inferior gyrus, and fusiform gyrus (FFG) were defined based on coordinates from prior literature indicating significant dynamic activity in face processing (Fox et al. 2009; Chan et al. 2011; Axelrod and Yovel 2013; Duchaine and Yovel 2015; Plank et al. 2021). The ROIs were selected and then labeled using the xjView toolbox (https://www.alivelearn.net/xjview). Moreover, masks from Neurosynth meta-analysis were also obtained for self-related masks using the keyword "self." ROI masks from AAL3 (Rolls et al. 2020) were constructed by resampling the AAL3 template using affine transformation. The PSC over these ROIs were calculated based on the work of Mazaika (2009) from the ArtRepair toolbox for SPM (https://cibsr.stanford.edu/tools/human-brain-project/artrepair-software.html) (Mazaika 2009). We tested our main hypothesis by calculating condition differences between PSC in the self-face processing ROIs and social brain ROIs.

Multivariate brain analysis

Apart from univariate analysis, multivariate analysis on the fMRI would result in similar but richer information on patterns of brain activation in different conditions (Pillet et al. 2020). Therefore, we performed the multivariate analysis to provide more detail about how OT influences brain activation would be revealed (see details below).

Principal component analysis

Because of the high dimensionality of fMRI data, we have conducted PCA for dimensionality reduction and identifying OT effects in task-based fMRI. Specifically, we performed whole-brain PCA on task fMRI activities and extracted the first fMRI component that explained the most variance in data. We then investigated importance of psychometric scores to the projection of this principal component (Supplementary Fig. S8). Principle component and projections were generated using *sklearn* package (Pedregosa et al. 2011). Reported principle components (Supplementary Fig. S8) were based on the activation pattern of all participants.

Multivariate pattern analysis

Unlike traditional analysis using univariate or mass-univariate approaches, the multivariate pattern analysis (MVPA) considers patterns of responses across multiple voxels, rather than single voxel-based or region-based values. The fMRI data are naturally multivariate, which allows for the multivariate analysis of multidimensional data. MVPA is a machine learning-based approach, mainly dealing with classification and regression problems. In this way, the activation of thousands of voxels in fMRI data is reduced to accuracies in several classifiers. In the current study, the analysis was conducted using the *Decoding Toolbox* (TDT) (Hebart et al. 2015). We first applied the boxcar function from the preprocessed fMRI data producing 4 beta estimates per run for each participant. Each of the 4 beta estimates was then correlated to the self-child, self-adult, other-child, and other-adult conditions, respectively. After the correlation, we did a group-level MVPA from the generated beta images. Under each 1 of the 4 conditions, we performed a whole-brain MVPA on discriminating OT or PL treatment with the searchlight method. The goal was to generate the accuracy of each voxel on discriminating the treatment.

To further examine the effect of OT on self-related processing, we also tested whether fMRI data of 2 groups in the present study can be reliably decoded from the ROIs involved in "self-referential," according to the Neurosynth meta-analysis. We first masked participants' brain data with the whole brain mask of self-referential-related regions from Neurosynth. The mask contains the statistically significant voxels from an automated meta-analysis of 166 studies, and the results from the association test were chosen (https://neurosynth.org/analyses/terms/self%20referential/).

Statistics for MVPA

The accuracy shown in Fig. 3 is masked by permuted significance with false discovery rate (FDR) correction (Benjamini–Hochberg method, Benjamini and Hochberg 1995). We used similar methods as demonstrated in previous literature (di Oleggio Castello et al. 2021). First, the actual accuracy of the classifiers was calculated with the correct labels assigned to each beta image. Then, we used 2 substeps to create a null distribution for the permutation testing. We started with repeatedly shuffling the labels assigned to the images and getting classifier accuracy based on the shuffled relationships between the labels and images. Note that the amount of the image in each condition stays the same after shuffling.

We repeated this substep 120 times for the whole-brain MVPA classifier and 200 times for the ROI masked MVPA classifier. The number of iterations is limited because of computational power limitations. After we have the shuffled MVPA classifier accuracy, we randomly choose *N* sets of classifier accuracy from the 120 (or 200) sets, where *N* equals the number of images used to train the classifier. The mean accuracy of the *N* sets was then calculated. This substep would be repeated 10,000 times. The result of this step would be a null distribution of classifier accuracy. Next, we could use the generated null distribution to make voxel-wise comparisons with the actual classifier accuracy. For each voxel, we count the number of accuracies higher than the actual accuracy and divide the total number by 10,000. The product of this procedure would give us the *P*-value of each voxel. The final step is to use FDR correction (Benjamini–Hochberg) and get the adjusted *P*-value. Figure 3 shows the accuracies of the voxels with *P*-values less than 0.05 after correction.

The significance of the difference between OT and PL classifiers is slightly different from above. The actual difference in accuracy is calculated with the accuracies of the OT classifier subtracting the accuracies of the PL classifier. After generating the null distributions for both treatments, we calculate the pair-wise difference of the 10,000 pairs of accuracies. In the counting substep, a difference is counted if the absolute value of it is greater than the absolute value of the actual difference.

Representational similarity analysis and representational connectivity analysis

To understand the similarity and difference between the 2 groups regarding the 4 facial conditions, we used the *NeuroRA* toolbox (Lu and Ku 2020) to extract representational dissimilarity matrices (RDM). We chose the same ROIs as used in the PSC analysis and calculated the RDM for each ROI from the OT or the PL group, respectively. We theorized that the self-reference-related regions would be influenced by OT. Therefore, we picked ROIs that had been previously reported to be associated with self-reference to run representational similarity analysis (RSA). Then, we were able to derive the correlation matrix among the ROIs. For comparison between the 2 correlation matrices, we used the Fisher *Z*-transformation to derive the significance of the OT and PL ROI connectivity difference (see Fig. 6).

Results

Psychometric data

The questionnaire scores of participants in the 2 treatment groups are summarized in Table 1. The means and standard deviations of the scores were calculated among each group, separated by treatments (OT and PL). We also performed pairwise comparison statistics between OT and PL treatment of these questionnaire scores (Table 1). The full names of each score are listed in the table caption. None of the 2-tailed *t*-tests were significant, providing no evidence for group differences between OT and PL treatment in participants' psychometric data. Furthermore, participants were asked to take each of the PA, NA, and SA tests twice, with one before the treatment administration and one after the treatment administration. The absence of significance changes provides no evidence regarding the role of OT on these trait or emotional state scores.

Behavioral performance

Supposing the correct answer for self-morphed faces is "yes" (see Section 2.2 for more details), we derived 4 conditions (namely, OT-child, OT-adult, PL-child, and PL-adult) and calculated the accuracy separately (Fig. 2a). We calculated overall accuracy for each participant and put into a 2-way mixed analysis of variance (ANOVA) using treatment groups (OT vs. PL) as between-subject factor and facial conditions (self vs. other, and child vs. adult face) as within-subject factors. The result of ANOVA showed a significant main effect of facial conditions (F (1, 57) = 54.6716, $P < 0.001$, $\eta_p^2 = 0.4949$), where participants showed higher accuracy on adult face discrimination (Fig. 2a). However, there is no significant effect of treatment (F (1, 57) = 0.1119, $P = 0.7392$, $\eta_p^2 = 0.0020$). These results are consistent with previous literature, which indicate that the participants could better detect faces at their own age (Wu et al. 2013). No other significant effects on accuracy were identified, *P*-values >0.05.

For RTs, similar analysis indicated that participants exhibited significantly longer RTs to discriminate self-morphed faces than other-morphed face (OT group: F (1, 29) = 18.6276, $P = 0.0002$, $\eta_p^2 = 0.3911$, PL group: F (1, 28) = 10.3001, $P = 0.0033$, $\eta_p^2 = 0.2689$). No significant effects on child/adult face on RTs were observed (Fig. 2b). To test whether there was RT difference for responding as "self" or "other" in 2 groups, we also conducted ANOVA on RTs with age (adult vs. child) by response type (responded as self vs. responded as other). It did not show significant treatment effect but suggested longer latency for responding as "self" for child faces than other types of faces (Supplementary Fig. S1).

As there were multiple conditions, overall accuracy would be an oversimplified generalization to the behavioral data. Confusion matrices for each participant were constructed, and we calculated the true positive rate and false positive rate from the confusion matrices. To better visualize the decisions of participants under each condition, we fitted the receiver operating characteristic (ROC) curve for each condition and calculated the detectability (d') and criteria (c) for each participant based on SDT. The ROC can illustrate the diagnostic ability for a specific condition, where we then compared the curves and their area under the curve (AUC) to see if there is any significant difference. Pair-wise significance comparisons showed that higher AUC for adult conditions than child conditions, indicating better performance to detect their

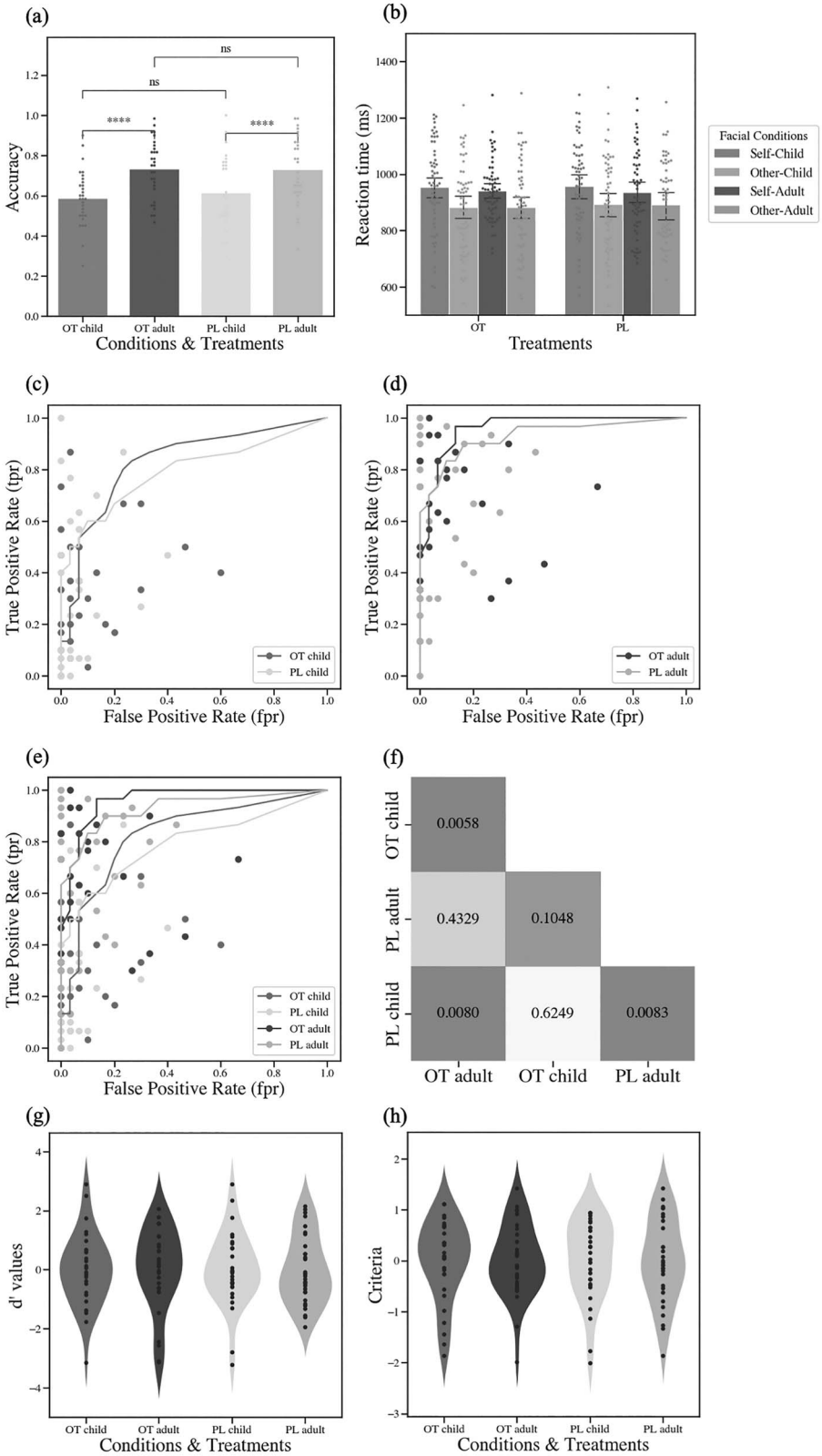

Fig. 2. Behavioral results of each condition. a) The mean accuracy under each condition for the 2 groups. The bars indicate the average accuracy for that condition, while the dots show accuracy of each participant in each run. The pair-wise comparisons are labeled in the figure, with "***" indicating ($1 * 10^{-4} < P < 1 * 10^{-3}$), and "ns" indicating ($0.05 < P < 0$) (nonsignificant). Neither of the between-group comparisons (OT vs. placebo) are significant, while both differences of the within group comparisons (child vs. adult face) are significant. The result of this *t*-test corresponds to the *F*-tests in ANOVA. b) The RTs of participants, bars are divided by treatment and facial conditions of the stimuli. c) Fitted ROC curves and *P*-values of the pair-wise comparisons. ROC curves for the child face conditions. d) ROC curves for the adult face conditions. e) Composite ROC plot for all 4 conditions. f) Heatmap of *P*-values for pair-wise ROC comparisons. *P*-values are labeled in the box for each corresponding comparisons. As indicated in the heatmap, the only significant comparisons are: OT child – OT adult, Placebo child – Placebo adult, and PL child – OT adult. This result is consistent with the ANOVA and *t*-test results, such that the behavioral results indicated significant adult versus child difference, but no treatment effect. g) The *d'* values of each condition. h) The response criteria of each condition. Values were calculated based on SDT analysis for 4 conditions in 2 groups. It indicated no significant effect on detectability and response criteria.

Table 2. Second-level univariate analysis table for OT versus PL contrast.

Conditions	Brain region	T-score	MNI coordinates		
			x	y	z
Self-child	Cuneus—L	10.01	−12	−103	4
	Cuneus—R	15.09	12	−103	1
	IFG—L	6.77	−30	23	−11
	IFG—R	7.94	36	20	−5
	FFG—L	12.38	−39	−58	−23
	FFG—R	13.11	45	−52	−20
Other-child	Cuneus—L	9.34	−12	−103	7
	Cuneus—R	12.17	15	−100	1
	IFG—L	5.07	−33	26	10
	IFG—R	4.96	36	23	−2
	FFG—L	7.61	−42	−55	−23
	FFG—R	9.97	42	−52	−20
Self-adult	Cuneus—L	10.92	−12	−103	4
	Cuneus—R	14.35	12	−103	4
	IFG—L	10.41	−33	20	−2
	IFG—R	9.80	33	23	−2
	FFG—L	12.22	−42	−55	−23
	FFG—R	14.04	42	−52	−20
Other-adult	Cuneus—L	11.90	−12	−103	4
	Cuneus—R	14.37	15	−100	1
	IFG—L	3.20	−42	17	−2
	IFG—R	7.96	48	35	16
	FFG—L	11.77	−42	−58	−23
	FFG—R	12.61	45	−52	−20

Note: Table showing regions with FDR-corrected $P < 0.05$ (Benjamini–Hochberg method). Coordinates were labeled using the xjView toolbox (https://www.alivelearn.net/xjview).

own faces in adult conditions (Fig. 2c–f). The ROC curve results were consistent with accuracy results, showing no significant effect on other factors. Furthermore, based on SDT and the work of Dal Martello and Maloney (2006), we computed the detectability and response criteria for each participant. However, as shown in the results of Fig. 2g and h, none of the comparisons was significant.

fMRI results

We combined multiple approaches in the fMRI data analysis, including univariate analysis, ROI analysis, MVPA, and RSA.

Univariate analysis

In an exploratory analysis, to identify activated clusters associated with group differences of face conditions, we conducted a group-level analysis on the OT versus PL contrasts. We observed significant activation differences (FDR-corrected $P < 0.05$, Benjamini–Hochberg) during face judgment between the OT and PL group, and the resulting statistics and coordination table are shown in Table 2. The cuneus (and occipital lobe in general) showed greater activity under all conditions for OT group, which indicated increased sensitivity to face stimuli under OT administration. Similar to previous fMRI findings (Platek et al. 2002; Platek et al. 2004), we found stronger IFG activity in self-morphed facial conditions compared to other-morphed facial conditions, further indicating the role of IFG in self–other distinction. Moreover, the FFG showed stronger activity on adult-face conditions compared to child-face conditions (Table 2).

Multivariate pattern analysis

We generated 2 MVPA classifiers to discriminate the 4 facial conditions based on images from OT and PL treatment groups. Using permutation methods (di Oleggio Castello et al. 2021), we calculated the significance of the classification accuracy on each voxel (Fig. 3a and b), and the significance of the differences between the OT and PL classifier (Fig. 3c). As shown in Fig. 3a, significant voxel accuracy for the OT classifier centers around regions such as IFG and visual processing areas, while the significant voxels were more scattered across the whole brain in the PL classifier (Fig. 3b). The whole brain difference result (Fig. 3c) clearly visualized this difference in pattern between the OT and PL classifier. It revealed that there are significant differences in areas such as occipital cortex, MFG, and IFG, suggesting the roles of visual perception brain area and IFG in self–other distinction. These MVPA findings also confirmed the univariate analysis results, showing the importance of the above-mentioned regions.

MVPA with ROIs from Neurosynth

To further examine the effect of OT on self-related processing, we tested whether fMRI data of 2 groups in the present study can be reliably decoded from

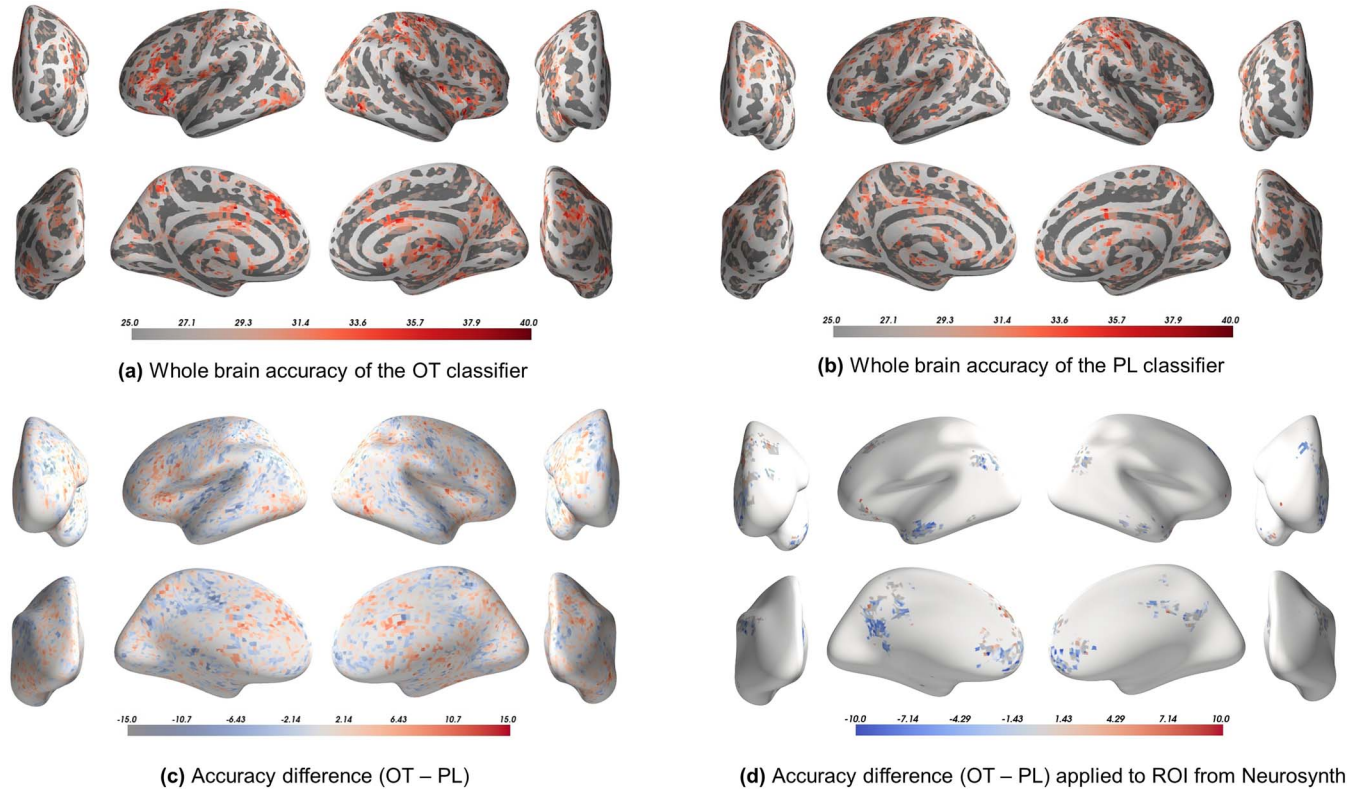

Fig. 3. MVPA classifier accuracy results. a–c) Whole brain MVPA classifier accuracy results, numbers displayed above the color bar are in percentage (%). Colored voxels were FDR-corrected (Benjamini–Hochberg) statistically significant voxels using the permutation method described in di Oleggio Castello et al. (2021). The chance level accuracy of the classifier discriminating the 4 facial conditions is 25%; therefore, the color bar has lower bound of 25%. MVPA results are consistent with the univariate results, showing significant differences between OT and PL classifier in areas such as IFG. d) MVPA results with masks from Neurosynth self-referential ROI. To classify self versus other conditions, there were only 2 conditions in the classifier; thus, the baseline accuracy of the classifiers is 50%. The color bar does not reflect this change, as we are only showing the difference plot. The gray regions represent the Neurosynth self-referential ROI.

the ROIs involved in "self-referential," according to the Neurosynth meta-analysis. We first masked participants' brain data with Neurosynth masks. The mask from Neurosynth is shown in Fig. 3d as the gray overlaid regions. Then, we trained classifiers on either OT or PL group data to discriminate between self (SC and SA) and other (OC and OA) conditions. Hence, the baseline accuracy for each voxel in both of the classifiers is 50%, compared to 25% in the whole-brain classifiers. Figure 3d shows the significant voxels on the difference between OT and PL condition in the Neurosynth ROI. The statistics were generated using a similar permutation method as specified in the above whole brain MVPA (Fig. 3c). This further implied that OT treatment might increase the activation difference between self and other conditions, resulting in the better performance in certain regions of the OT classifier.

ROI results from AAL3

PSC values for each of the 16 selected ROIs were calculated under each condition to identify the difference between self and other faces for 2 groups. For plots of other selected ROIs and detailed statistic information, see Supplementary Figs. S2–S4. Figure 4 shows the PSC plots of the bilateral ACC, IFG-operant part, IFG-triangular part, and insula. It showed that OT increased self–other differentiation, particularly for adult faces, which effect was more pronounced in left hemisphere regions. With self/other face and child/adult face as 2 factors, the 2-way ANOVA on PSC of IFG revealed significant influence from self/other face. For instance, the FDR-corrected *P*-values of right IFG (opercular part) on the self/other factor were OT: $P < 0.0001$, PL: $P = 0.0032$. However, on the child/adult factor, the corrected *P*-values were both nonsignificant (FDR-corrected $P = 1$). Additionally, we also observed that for the interaction term, the OT group has significantly corrected *P*-value, while the PL group does not (OT: $P = 0.0144$, PL: $P = 0.2117$). Other regions showed similar trend include left IFG (triangular part) (OT: $P = 0.0192$, PL: $P = 0.2163$) and right IFG (triangular part) (OT: $P = 0.0139$, PL: $P = 0.1768$). This similar trend might indicate that administration of OT is correlated with the IFG region using information from both child/adult and self/other factors. In our Supplementary Material (Supplementary Fig. S4), we showed the significance of all ROIs. *P*-values shown in Supplementary Fig. S4 and presented in the current section are all adjusted by FDR correction (Benjamini–Hochberg method, Benjamini and Hochberg 1995). Other ROIs showed the similar pattern, especially over the left hemisphere.

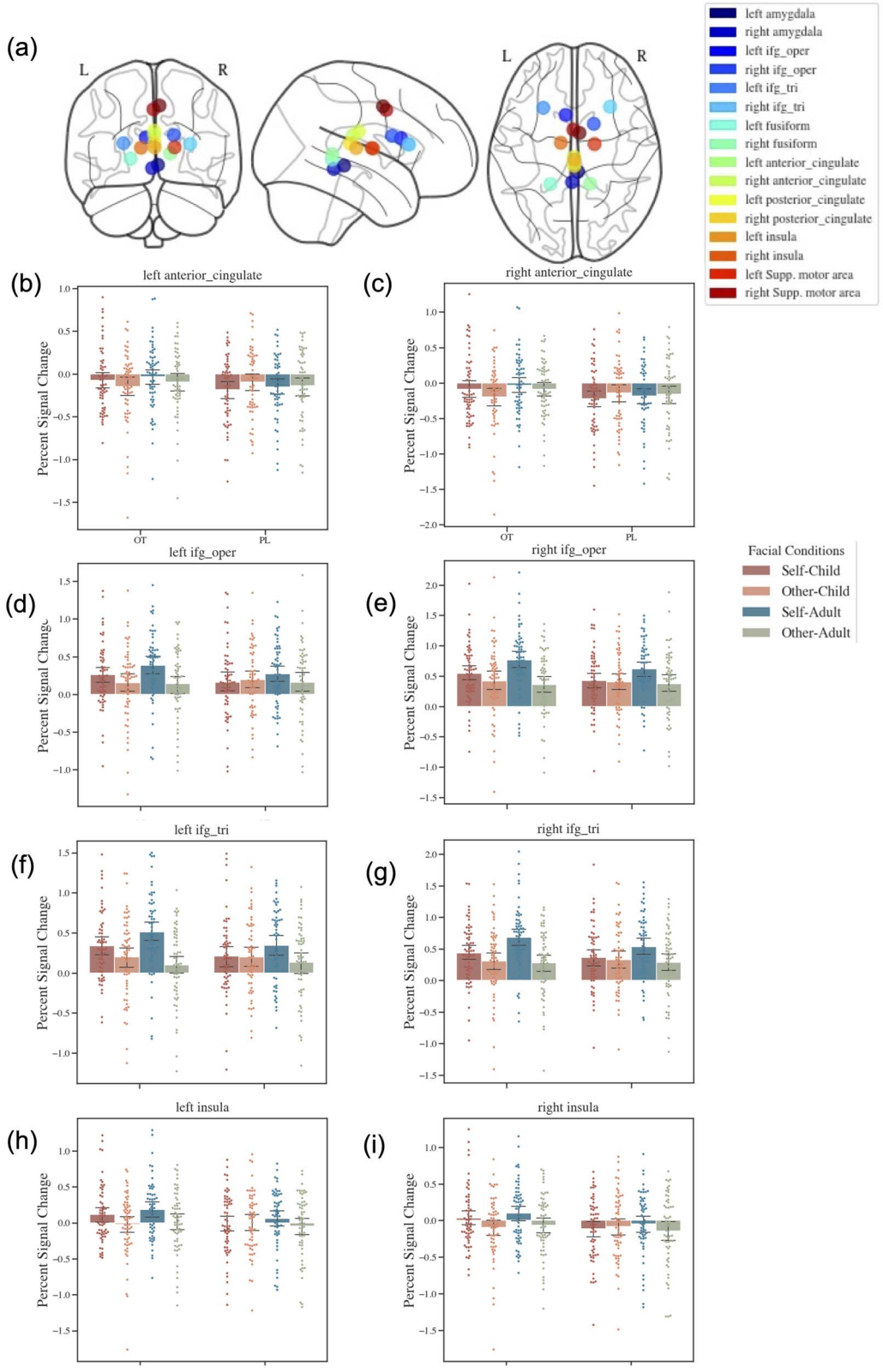

Fig. 4. a) Visualization of center of mass locations for each selected ROIs from the AAL3 template. Note that this is a rough representation of the location of the ROIs, calculated by averaging coordinates in each of the ROIs. b–i) Percent Signal Change over IFG (in AAL3 template) of the ROI analyzed. ROI names are noted in the titles of the subfigures. Error bars show the bootstrapped confidence intervals with 500 iterations. The PSC plots (b–i) share the same facial conditions legend.

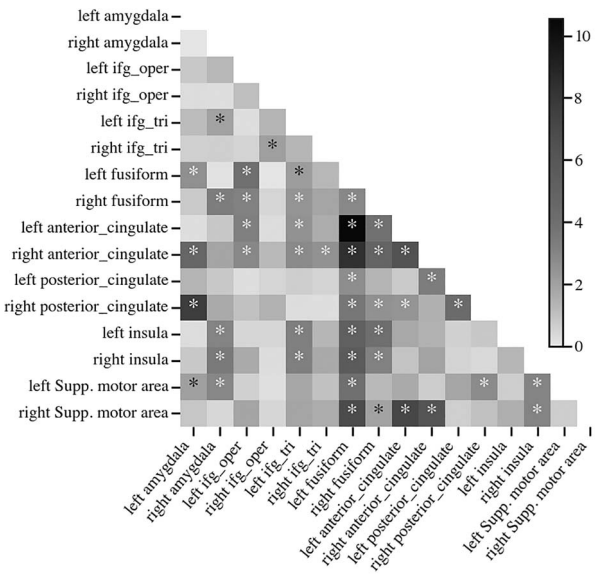

Fig. 5. Z-scores of connectivity-similarity difference between OT and PL group. Statistics for correlation difference between RDM correlations of OT and PL, using Fisher Z-transformation. Labels on the axis are the names of the selected ROIs, the same as they appeared in the AAL3 atlas. An asterisk ("∗") indicates significance in that pair of correlation difference.

Representational connectivity difference between OT and PL group

To explore the possible correlations between activation of the prior ROIs, we calculated the similarity of the activation for each ROI in both OT and PL groups, respectively. With Fisher Z-transformation, statistics for correlation difference between RDM correlation of OT and PL were calculated. The significance of OT - PL difference is shown in Fig. 5. There was ROI connectivity difference (e.g. amygdala and ACC, FFG, and insula) between groups, suggesting a higher representational connectivity over OT group in these face and self-referential ROIs.

Correlations among behavioral data, psychometric data, and fMRI data

Up to this point, we have displayed the results of 3 types of data: psychometric data, behavioral data, and fMRI data. We were able to derive different implications from these data of different modalities. In this section, we will bridge the data and provide interpretations of their relationships.

Correlation between ROI activation

Figure 6 shows the difference of OT groups' and PL groups' correlation among ROI activities. The ROI–ROI correlations indicated the co-activation pattern in the self and face perception brain regions. The figure showed that more pairs of ROIs were significant in the adult face conditions ("self-adult" and "other-adult"). This indicates that compared to child-morphed stimuli, OT might modulate adult-morphed faces and result in changing co-activation pattern in face-related and self-related ROI.

Correlation between psychometric data and fMRI data

Since we have found interesting OT effects on self–other differentiation based on fMRI data, to investigate the relationship between brain activities and psychometric scores in participants, we conducted bivariate correlation analysis between behavioral scores and PSC activity difference (self–other) over selected ROIs (similar as difference calculation in the ROI–ROI analysis). We first calculated the correlations between behavioral data and neural activation of self versus other condition, respectively. Based on experimental design and previous analysis on psychometric data, OT and PL group difference was not considered. Then, similar to previous processes, we derived the difference between self and other and calculated the significance. Figure 7 shows the z-scores of the difference. The only significant difference appeared in child condition: the correlation difference between left insula and Machiavellianism Scale (from psychometric data) was significant. This is consistent with previous research, where individuals with different Machiavellianism Scale scores exhibit different volume and activation in the insula areas (Verbeke et al. 2011; Deak et al. 2017).

Although administering different treatments has little influence on psychometric data, we have shown that treatment can have influence on fMRI activations. Therefore, the correlation patterns of psychometric data and ROI activation between OT and PL treatment could be different. As shown in Fig. 8, the difference of correlation between OT and PL group was more significant in other-face conditions (other-child and other-adult) compared to self-face conditions. Specifically, the IRI showed more significant difference in other-face conditions. We found that in OT participants, IRI have higher correlations with the ROI activations (see Fig. 8). This trend showed not only in other-face conditions but also in self-face conditions, albeit less significant. A possible explanation is that OT modifies the activation in the selected brain areas and consequently their correlation with the IRI scores. It is possible that brain regions of participants under administration of OT would activate more differently when seeing an other-morphed face, resulting in a more significant difference.

Discussion

Accumulating evidence indicates that the OT influences human behavior (Zhu et al. 2019) and brain activity (Wu, Feng, et al. 2020a; Zheng, Liang, et al. 2021a; Zheng, Punia, et al. 2021b). Since the previous behavioral results were conflicted on whether OT sharpens or blurs the self–other distinction (Colonnello et al. 2013; Yue et al. 2020; Zhao et al. 2020), we investigated the effect of OT on self–other differentiation with self-morphed adult and child faces using comprehensive behavioral and neural analyses. Although we did not observe a significant effect of OT on behavioral responses, we were able to show the following: (i) OT increases brain activity in self- and

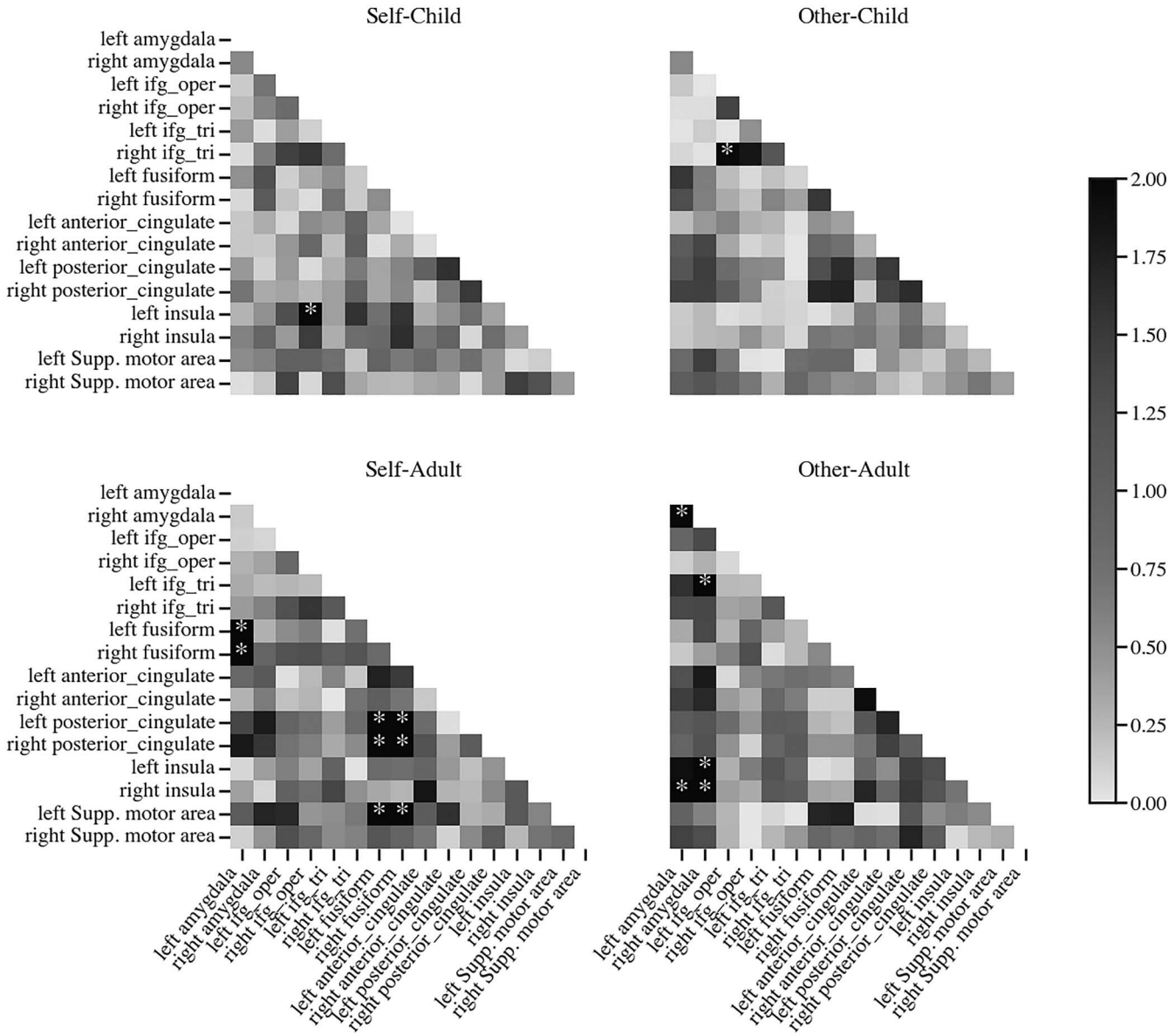

Fig. 6. Significance of correlation differences between OT and PL group, divided by the 4 facial conditions. Significance was calculated using the Fisher Z-transformation. Asterisk (*) indicates significance.

face-related brain regions, specifically in the IFG and visual areas; (ii) the voxels over the visual cortex for self-morphed child faces can classify the OT versus PL group; and (iii) the OT versus PL self–other differentiation effect in brain activity is more pronounced for adult faces in the left hemisphere.

Effects of OT on self-relevant face processing at the neural level but not on behavioral performance

Regarding behavioral performance, a previous study indicates that OT reduced RTs for making both self and other judgments (Colonnello et al. 2013) and reduced the recall accuracy (Zhao et al. 2016). However, we did not observe a significant OT modulation effect on accuracy (Fig. 2a), RTs (Fig. 2b), or detectability in SDT analysis (Fig. 2c–f) in the current self–other differentiation task. Although nonsignificant, our behavioral analysis showed that the OT group had better diagnostic ability than the PL group (Fig. 2e). OT could possibly facilitate overall behavioral performance in the current task. Moreover, in the behavioral results, we observed significantly better performance for adult-morphed facial stimuli (Fig. 2c). This significance might be attributed to the dominant age effect, that adult individuals would have better performance toward adult facial stimuli, along with generally higher accuracy and shorter RTs.

Although there were no significant differences, we observed that the OT group had a relatively better performance than the PL group, especially in the ROC results (Fig. 2e). However, previous studies have reported similar and significant results (Colonnello et al. 2013; Zhao et al. 2016). This prompted us to examine the brain activity evidence, as participants' behavioral performance can be

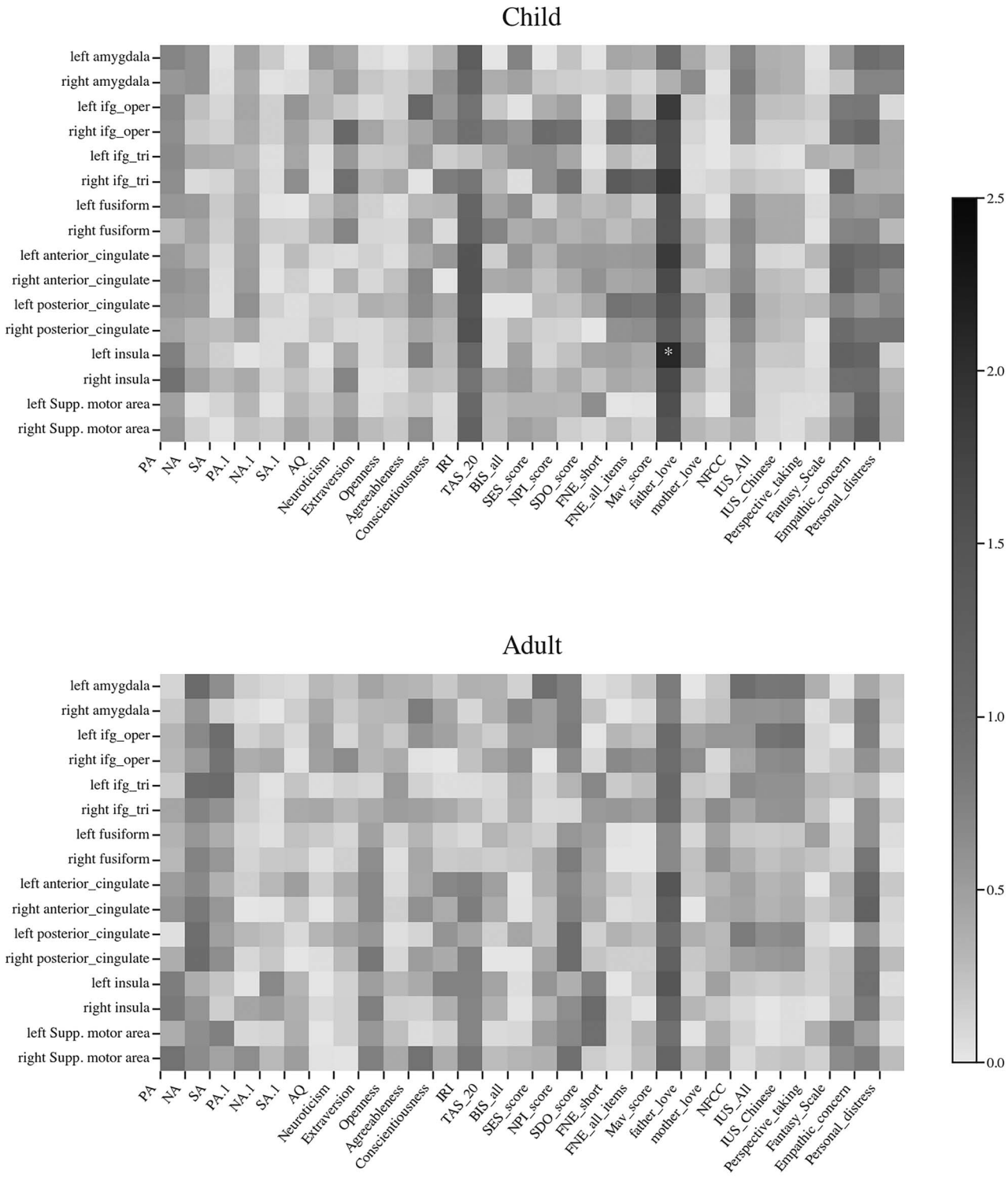

Fig. 7. Significance of correlation differences between self condition and other condition, divided by child-face and adult-face condition. Asterisk (*) indicates significance.

influenced by various factors, albeit in a well-controlled environment and carefully planned experiments. Therefore, we calculated the differences in fMRI data using a comprehensive set of analyses. The goal was to determine whether there were differences despite the non-significance of the behavioral results. In our fMRI results, we observed a general increase in brain activation over the visual area (cuneus) and IFGs (Table 2) in the OT group compared with that in the PL group. As proposed in previous studies, facial features are first encoded by the visual area, followed by the encoding of self-referential properties by the frontal cortex (Prince et al. 2009). Therefore, our second-level result indicating brain regions in processing faces within the occipital cortex (FFG and

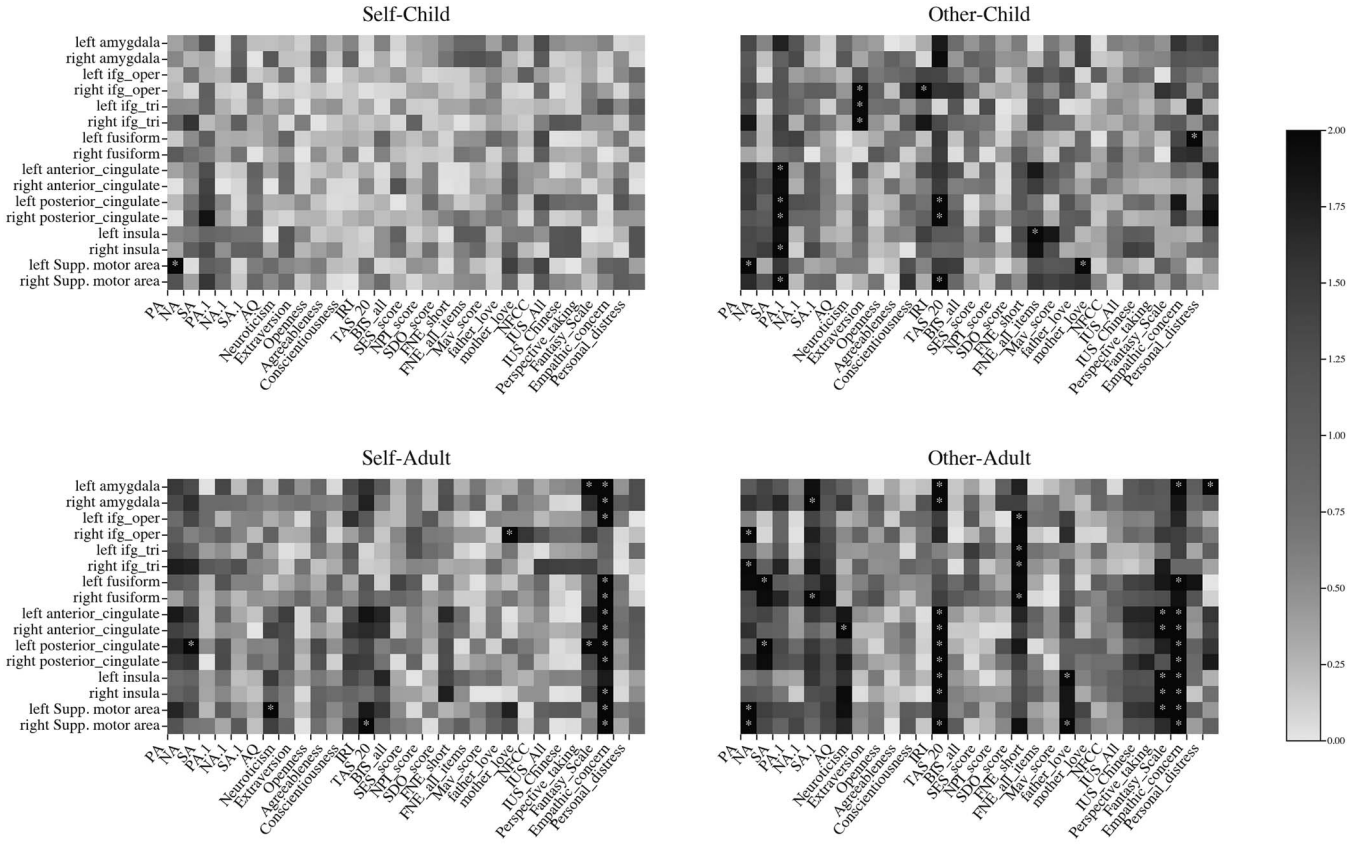

Fig. 8. Significance of correlation differences between OT and PL groups. Asterisk (*) indicates significance. Separate correlation heatmap of OT and PL groups is shown in Supplementary Fig. S7.

cuneus) was consistent with the processing of faces with regions within the FFG (Puce et al. 1995; Kanwisher et al. 1997; Rossion et al. 2003). According to Uddin et al. (2005), the IFG in the right hemisphere responded parametrically to the amount that a facial stimulus looked like the participant (Uddin et al. 2005); the increased IFG activity for the OT group in our study suggests more self-resemblance processing after OT administration. The ROI analysis and MVPA results further confirmed that there were significant differences in the brain regions, such as the IFG (Fig. 3c). With the self-referential masks, we also found differences between the classifiers when discriminating self or other facial stimuli (Fig. 3d), which provides evidence that OT administration modulates the self–other distinction in the brain.

For the ROIs in face processing and self-referential processing, we further observed a greater self–other distinction in ROI activity, especially in adult-morphed faces. For example, we found that the activities of the IFG and insula are greater for self-adults than for other-adults. The inferior parietal cortex, along with the prefrontal cortex, comprises a self–other brain network, which is important in distinguishing the self from the other, and shared self–other discrimination is considered a root for prosociality, such as empathy and trust (Decety and Sommerville 2003). Interestingly, previous studies have shown that viewing the face activated the IFG, inferior parietal lobe, and inferior occipital cortex, especially in the right hemisphere (Sugiura et al. 2005; Uddin et al. 2005; Kaplan et al. 2008). However, in our study, the OT group showed a stronger self–other distinction in the left hemisphere. Supposing there is a right hemisphere advantage in the "self" network, we provided the first evidence of OT sharpening the left hemisphere self-network, supporting greater self–other distinction. Thus, the higher self–other distinction of our ROI might reflect the pronounced role of the left hemisphere area, which may be responsible for maintaining self–other distinctions under OT administration. This may also reflect a flexible adaptation to update the self–other distinction after OT administration.

Consistent with this possibility, the MVPA results provide further evidence for our explanation. It first confirmed that face processing and self-referential processing of brain regions can be used to classify the 2 treatment groups. Furthermore, the results indicated that OT could enhance the classification between self-morphed and other faces in the voxels of self-referential ROIs. Given that OT increased self–other distinction in the left hemisphere, better classification performance in self ROIs seems more plausible. Supporting the higher social salience or flexibility in the OT group, it demonstrated further evidence relevant to social cognition that such flexibility may be due to hemispheric balance.

Child face versus adult face: higher OT effect on adult faces

It is notable that the self–other distinction was stronger in adult-morphed faces than in child-morphed faces, regardless of the treatment conditions (OT and PL) (Fig. 2a, Fig. 2c–f). This overall adult face advantage could reflect an own-age bias (Wu et al. 2013), such that the recognition of faces of one's own age group is often better than that of another age group (Rhodes and Anastasi 2012). Accordingly, with respect to the age cohort of participants (adults), the higher performance of adult faces across the 2 groups reflected a typical advantage in self-age face processing. Research has also shown that, compared to out-group faces, in-group faces receive more holistic and in-depth processing, which facilitates the perceptual discrimination of faces (Sporer 2001; Hugenberg and Corneille 2009). The participants of the current study were all adults, hence showing an in-group tendency toward adult faces. Our results are consistent with these findings.

Another possible explanation for these results is that people have various capacities to recognize their own current and past facial appearance (Apps et al. 2012). Studies have suggested that different brain regions process one's current and childhood faces (Apps et al. 2012). The inferior occipital gyrus, superior parietal lobule, and inferior temporal gyrus are more involved in current self-processing, while the TPJ and inferior parietal lobule are more involved in childhood self-processing. Accordingly, OT may only affect self–other differentiation in recognition of the current self-face, but not the past self-face, as the latter requires more memory encoding and retrieval (Fink et al. 1996).

Therefore, the better detection accuracy in the adult face condition is consistent with our previous work (Wu et al. 2013) and may reflect the automatic processing of self-age faces. For adult participants, a child-morphed stimulus may be required for the retrieval and maintenance of childhood self-images because the image is not their current facial appearance. Thus, there is a more indirect link between childhood image and the response "self" than adult-morphed faces and "self." This would lead to lower accuracy, longer RTs, and no OT enhancement effect on brain activity. In this vein, how OT modulates self-face recognition across ages remains an open question, but the current different effects on self–other distinction in adult and child faces may yield important insights.

Individual differences of the OT effect: brain–behavior association

To test whether the effect of OT is associated with individual differences in personal traits, we collected psychometric data from the participants. Psychometric data were collected to control and observe the homogeneity of the 2 groups (OT and PL). Following previous work showing individual differences on the effect of OT, our findings also offer novel associations between brain activity, functional connectivity, and personality traits. A recent study suggested that the OT effect on self–other distinction is influenced by the OT receptor genotype (Zhao et al. 2020), as OT only modulates the task decision time of rs53576G carriers. It is possible that there is a similar link between individual differences and the effects of OT on our findings. In the analysis of psychometric data, we found correlations between the participants' psychometric questionnaire scores and functional connectivity in the fMRI task, indicating that the effect of OT on the self-brain might be modulated by individual differences.

For example, at the correlation level, we observed a significant brain–behavior correlation (IFG activity and measured IRI score) difference. This means that the OT group has a higher IRI score and IFG activity correlation for the other-child condition than the PL group. Since self–other overlap is critical for empathy, studies have shown a correlation between personal empathy traits and self–other overlap (Batson et al. 1997; Oveis et al. 2010; Cooke et al. 2018). To some extent, OT strengthens the link between individual differences in social adaptation (Ma et al. 2016). That is, after OT administration, people with higher IRI might enhance the self–other overlap, while people with lower IRI might enhance the self–other distinction. Together with other correlations, these findings support the view that differential pharmacological effects are rooted in individual differences in personality traits.

Limitations and future perspectives

This study had several limitations that should be addressed. First, the current study has a between-subject design, which is unable to directly track the neural response change before and after OT administration. Second, the sample size of the participants and the male-only participant group may not allow the generalization of the conclusions. Previous literature has shown that males are less likely to be influenced by hormonal fluctuations (Kimura and Hampson 1994; Teicher et al. 1995; Sisk and Foster 2004; Winslow and Insel 2004; Sisk and Zehr 2005). Additionally, our study used a self–other facial discrimination task, which has been proven to show male advantage in the results (Wu et al. 2013). Therefore, we recruited male-only participants. However, since there are gender differences in the effects of stress and OT on the self–other distinction (Tomova et al. 2014), future research could utilize a larger sample size with more diverse participants and likely capture gender differences in the effect of OT. Further research could also combine pre- and post-administration tests to investigate large-scale neural networks and to track and reveal the underlying neural mechanism of the OT effect in self–other distinction processing.

The present findings substantially extend previous findings on the effect of OT on neural responses to social

salient stimuli (Wittfoth-Schardt et al. 2012). Our study first used self-morphed faces to investigate the OT effect on self-relevant processing, more specifically, to tap into the self–other distinction. Evidence has shown distinct neural underpinnings of the processes of self- and other-related information, which are critical for human social motivation and behaviors (Lieberman 2007). A previous study using trait judgment tasks showed the effects of OT on self-referential processing, including reduced RTs for self-related trait judgments, increased accuracy in memory retrieval, and decreased MPFC activation for self-related trait adjectives (Liu et al. 2017). Although our findings showed no significant effects of OT on behavioral judgments and RTs, the comprehensive fMRI analysis results indicated the robustness of increased self–other distinction in the OT group, especially for adult faces.

Conclusions

The current findings extend previous research linking the effect of OT with self-relevant information processing, providing potential core mechanisms associated with OT modulation effects on social behaviors. Our results highlight significantly larger self–other differences in brain activity, particularly within the left face and self-referential processing brain regions. This finding suggests a possible OT-induced complementary left self-brain network. These OT effects were associated with personality traits, which confirmed the individual differences in OT effects of self–other distinction brain activity. This may shed light on the role of self-relevant processing in social behavioral changes under OT administration.

Author contributions

HW designed and performed research. YW and HW analyzed data. YW, RW, and HW wrote the paper.

Acknowledgments

The authors would like to thank Qingyuan Wu, Shensheng Wang, Yanghua Ye, Quanying Liu, and Kun Chen for comments on the early version of the manuscript. We would like to thank Naomi Hein from Harvard Medical School for her help on proof reading. We also thank all research assistants who provided general support in participants recruiting and data collection.

Supplementary material

Supplementary material is available at *Cerebral Cortex* online.

Funding

This study is supported by the Science and Technology Development Fund (FDCT) of Macau (0127/2020/A3), National Natural Science Foundation of China (U1736125), Natural Science Foundation of Guangdong Province (2021A1515012509), Shenzhen-HongKong-Macao Science and Technology Innovation Project (SGDX20201103092801 00), and SRG of University of Macau (SRG2020-00027-ICI).

Conflict of interest statement: The authors declare that there is no conflict of interest in the current study.

Data and Code availability

The data are available upon request to the corresponding author. The code for data analysis is available at https://github.com/andlab-um/OT face.

Ethical statement

The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All procedures performed in this study involving human participants were in accordance with the Declaration of Helsinki (as revised in 2013). The ethics committee of Beijing Normal University approved this study and written informed consents were obtained from all participants.

References

Adolphs R. The neurobiology of social cognition. *Curr Opin Neurobiol*. 2001:11:231–239.

Adolphs R. Cognitive neuroscience of human social behaviour. *Nat Rev Neurosci*. 2003:4:165–178.

Adolphs R. The social brain: neural basis of social knowledge. *Annu Rev Psychol*. 2009:60:693–716.

Adolphs R. Conceptual challenges and directions for social neuroscience. *Neuron*. 2010:65:752–767.

Alvergne A, Faurie C, Raymond M. Differential facial resemblance of young children to their parents: who do children look like more? *Evol Hum Behav*. 2007:28:135–144.

Amodio DM, Frith CD. Meeting of minds: the medial frontal cortex and social cognition. *Nat Rev Neurosci*. 2006:7:268–277.

Apps MA, Tajadura-Jiménez A, Turley G, Tsakiris M. The different faces of one's self: an fMRI study into the recognition of current and past self-facial appearances. *NeuroImage*. 2012:63:1720–1729.

Axelrod V, Yovel G. The challenge of localizing the anterior temporal face area: a possible solution. *NeuroImage*. 2013:81:371–380.

Barrett LF, Satpute AB. Large-scale brain networks in affective and social neuroscience: towards an integrative functional architecture of the brain. *Curr Opin Neurobiol*. 2013:23:361–372.

Bate S, Cook SJ, Duchaine B, Tree JJ, Burns EJ, Hodgson TL. Intranasal inhalation of oxytocin improves face processing in developmental prosopagnosia. *Cortex*. 2014:50:55–63. https://doi.org/10.1016/j.cortex.2013.08.006. https://www.sciencedirect.com/science/article/pii/S0010945213002086.

Bate S, Bennetts R, Parris BA, Bindemann M, Udale R, Bussunt A. Oxytocin increases bias, but not accuracy, in face recognition line-ups. *Soc Cogn Affect Neurosci*. 2015:10:1010–1014. https://doi.org/10.1093/scan/nsu150. https://pubmed.ncbi.nlm.nih.gov/25433464.

Batson CD, Sager K, Garst E, Kang M, Rubchinsky K, Dawson K. Is empathy induced helping due to self–other merging? *J Pers Soc Psychol*. 1997:73:495–509.

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

Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B Methodol*. 1995:57:289–300.

Brédarta S, French RM. Do babies resemble their fathers more than their mothers? A failure to replicate Christenfeld and Hill (1995). *Evol Hum Behav*. 1999:20:129–135.

Bressan P, Dal Martello MF. Talis pater, talis filius: perceived resemblance and the belief in genetic relatedness. *Psychol Sci*. 2002:13:213–218.

Chan AW, Downing PE et al. Faces and eyes in human lateral prefrontal cortex. *Front Hum Neurosci*. 2011:5:51.

Colonnello V, Chen FS, Panksepp J, Heinrichs M. Oxytocin sharpens self-other perceptual boundary. *Psychoneuroendocrinology*. 2013:38:2996–3002.

Cooke AN, Bazzini DG, Curtin LA, Emery LJ. Empathic understanding: benefits of perspective-taking and facial mimicry instructions are mediated by self-other overlap. *Motiv Emot*. 2018:42:446–457.

Dal Martello MF, Maloney LT. Where are kin recognition signals in the human face? *J Vis*. 2006:6:1356–1366.

Davis MH. Measuring individual differences in empathy: evidence for a multidimensional approach. *J Pers Soc Psychol*. 1983:44:113–126.

Deak A, Bodrogi B, Biro B, Perlaki G, Orsi G, Bereczkei T. Machiavellian emotion regulation in a cognitive reappraisal task: an fMRI study. *Cogn Affect Behav Neurosci*. 2017:17:528–541.

DeBruine LM. Facial resemblance enhances trust. *Proc R Soc Lond Ser B Biol Sci*. 2002:269:1307–1312.

Decety J, Sommerville JA. Shared representations between self and other: a social cognitive neuroscience view. *Trends Cogn Sci*. 2003:7:527–533.

Denny BT, Kober H, Wager TD, Ochsner KN. A meta-analysis of functional neuroimaging studies of self-and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *J Cogn Neurosci*. 2012:24:1742–1752.

di Oleggio Castello MV, Haxby JV, Gobbini MI. Shared neural codes for visual and semantic information about familiar faces in a common representational space. *Proc Natl Acad Sci*. 2021:118.

Domes G, Heinrichs M, Gläscher J, Büchel C, Braus DF, Herpertz SC. Oxytocin attenuates amygdala responses to emotional faces regardless of valence. *Biol Psychiatry*. 2007:62:1187–1190.

Duchaine B, Yovel G. A revised neural framework for face processing. *Annu Rev Vision Sci*. 2015:1:393–416.

Fink GR, Markowitsch HJ, Reinkemeier M, Bruckbauer T, Kessler J, Heiss W-D. Cerebral representation of one's own past: neural networks involved in autobiographical memory. *J Neurosci*. 1996:16:4275–4282.

Fox CJ, Iaria G, Barton JJ. Defining the face processing network: optimization of the functional localizer in fMRI. *Hum Brain Mapp*. 2009:30:1637–1651.

Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RS. Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp*. 1994:2:189–210.

Frith CD. The social brain? *Philos Trans R Soc B Biol Sci*. 2007:362:671–678.

Frith CD, Frith U. The neural basis of mentalizing. *Neuron*. 2006:50:531–534.

Gallagher HL, Frith CD. Functional imaging of 'theory of mind'. *Trends Cogn Sci*. 2003:7:77–83.

Gobbini MI, Gentili C, Ricciardi E, Bellucci C, Salvini P, Laschi C, Guazzelli M, Pietrini P. Distinct neural systems involved in agency and animacy detection. *J Cogn Neurosci*. 2011:23:1911–1920.

Hebart MN, Görgen K, Haynes J-D. The decoding toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. *Front Neuroinform*. 2015:8:88.

Hugenberg K, Corneille O. Holistic processing is tuned for in-group faces. *Cogn Sci*. 2009:33:1173–1181. https://doi.org/10.1111/j.1551-6709.2009.01048.x.

Hurlemann R, Patin A, Onur OA, Cohen MX, Baumgartner T, Metzler S, Dziobek I, Gallinat J, Wagner M, Maier W et al. Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. *J Neurosci*. 2010:30:4999–5007.

Kanwisher N, McDermott J, Chun MM. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci*. 1997:17:4302–4311.

Kaplan JT, Aziz-Zadeh L, Uddin LQ, Iacoboni M. The self across the senses: an fMRI study of self-face and self-voice recognition. *Soc Cogn Affect Neurosci*. 2008:3:218–223.

Keenan JP, Wheeler MA, Gallup GG Jr, Pascual-Leone A. Self-recognition and the right prefrontal cortex. *Trends Cogn Sci*. 2000:4:338–344.

Kelley WM, Macrae CN, Wyland CL, Caglar S, Inati S, Heatherton TF. Finding the self? An event-related fMRI study. *J Cogn Neurosci*. 2002:14:785–794.

Kimura D, Hampson E. Cognitive pattern in men and women is influenced by fluctuations in sex hormones. *Curr Dir Psychol Sci*. 1994:3:57–61.

Krienen FM, Tu P-C, Buckner RL. Clan mentality: evidence that the medial prefrontal cortex responds to close others. *J Neurosci*. 2010:30:13906–13915.

Lamm C, Bukowski H, Silani G. From shared to distinct self–other representations in empathy: evidence from neurotypical function and socio-cognitive disorders. *Philos Trans R Soc B Biol Sci*. 2016:371:20150083.

Lieberman MD. Social cognitive neuroscience: a review of core processes. *Annu Rev Psychol*. 2007:58:259–289.

Liu Y, Wu B, Wang X, Li W, Zhang T, Wu X, Han S. Oxytocin effects on self- referential processing: behavioral and neuroimaging evidence. *Social Cogn Affect Neurosci*. 2017:12:1845–1858.

Lu Z, Ku Y. Neurora: a python toolbox of representational analysis from multi-modal neural data. *Frontiers in neuroinformatics*. 2020:61.

Ma Y, Shamay-Tsoory S, Han S, Zink CF. Oxytocin and social adaptation: insights from neuroimaging studies of healthy and clinical populations. *Trends Cogn Sci*. 2016:20:133–145.

Martin AK, Huang J, Hunold A, Meinzer M. Dissociable roles within the social brain for self–other processing: a HD-TDCS study. *Cereb Cortex*. 2019:29:3642–3654.

Marusak HA, Carré JM, Thomason ME. The stimuli drive the response: an fMRI study of youth processing adult or child emotional face stimuli. *NeuroImage*. 2013:83:679–689.

Mazaika, P, Hoeft F, Glover, GH, Reiss, AL. Methods and Software for fMRI Analysis for Clinical Subjects. (2009). https://cibsr.stanford.edu/tools/human-brain-project/artrepair-software.html#documentation.

Mitchell JP, Macrae CN, Banaji MR. Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*. 2006:50:655–663.

Oveis C, Horberg EJ, Keltner D. Compassion, pride, and social intuitions of self-other similarity. *J Pers Soc Psychol*. 2010:98:618.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Pretten-Hofer P, Weiss R, Dubourg V et al. Scikit-learn: machine learning in Python. *J Mach Learn Res*. 2011:12: 2825–2830.

Pillet I, Op de Beeck H, Lee Masson H. A comparison of functional networks derived from representational similarity, functional connectivity, and univariate analyses. *Front Neurosci*. 2020:13:1348.

Plank IS, Hindi Attar C, Kunas SL, Bermpohl F, Dziobek I. Increased child-evoked activation in the precuneus during facial affect recognition in mothers. *Hum Brain Mapp*. 2021.

Platek SM, Kemp SM. Is family special to the brain? An event-related fMRI study of familiar, familial, and self-face recognition. *Neuropsychologia*. 2009:47:849–858.

Platek SM, Burch RL, Panyavin IS, Wasserman BH, Gallup GG Jr. Reactions to children's faces: resemblance affects males more than females. *Evol Hum Behav*. 2002:23:159–166.

Platek SM, Raines DM, Gallup GG Jr, Mohamed FB, Thomson JW, Myers TE, Panyavin IS, Levin SL, Davis JA, Fonteyn LC et al. Reactions to children's faces: males are more affected by resemblance than females are, and so are their brains. *Evol Hum Behav*. 2004:25: 394–405.

Platek SM, Mohamed FB, Gallup GG Jr. Contagious yawning and the brain. *Cogn Brain Res*. 2005:23:448–452.

Prince SE, Dennis NA, Cabeza R. Encoding and retrieving faces and places: distinguishing process-and stimulus-specific differences in brain activity. *Neuropsychologia*. 2009:47:2282–2289.

Puce A, Allison T, Gore JC, McCarthy G. Face-sensitive regions in human extrastriate cortex studied by functional MRI. *J Neurophysiol*. 1995:74:1192–1199.

Rhodes MG, Anastasi JS. The own-age bias in face recognition: a meta-analytic and theoretical review. *Psychol Bull*. 2012:138: 146–174. doi: https://doi.org/10.1037/a0025750.

Rimmele U, Hediger K, Heinrichs M, Klaver P. Oxytocin makes a face in memory familiar. *J Neurosci*. 2009:29:38. https://doi.org/10.1523/JNEUROSCI.4260-08.2009. http://www.jneurosci.org/content/29/1/38.abstract.

Rolls ET, Huang C-C, Lin C-P, Feng J, Joliot M. Automated anatomical labelling atlas 3. *NeuroImage*. 2020:206:116189.

Rossion B, Schiltz C, Crommelinck M. The functionally defined right occipital and fusiform "face areas" discriminate novel from visually familiar faces. *NeuroImage*. 2003:19:877–883.

Shamay-Tsoory SG, Abu-Akel A. The social salience hypothesis of oxytocin. *Biological Psychiatry* 2016:79:194–202. https://doi.org/10.1016/j.biopsych.2015.07.020.

Sisk CL, Foster DL. The neural basis of puberty and adolescence. *Nat Neurosci*. 2004:7:1040–1047.

Sisk CL, Zehr JL. Pubertal hormones organize the adolescent brain and behavior. *Front Neuroendocrinol*. 2005:26:163–174.

Sporer SL. Recognizing faces of other ethnic groups: an integration of theories. *Psychol Public Policy Law*. 2001:7:36–97. https://doi.org/10.1037/1076-8971.7.1.36.

Stanley DA, Adolphs R. Toward a neural basis for social behavior. *Neuron*. 2013:80:816–826.

Sugiura M, Watanabe J, Maeda Y, Matsue Y, Fukuda H, Kawashima R. Cortical mechanisms of visual self-recognition. *NeuroImage*. 2005:24:143–149.

Teicher MH, Andersen SL, Hostetter JC Jr. Evidence for dopamine receptor pruning between adolescence and adulthood in striatum but not nucleus accumbens. *Dev Brain Res*. 1995:89:167–172.

Tillman R, Gordon I, Naples A, Rolison M, Leckman JF, Feldman R, Pelphrey KA, McPartland JC. Oxytocin enhances the neural efficiency of social perception. *Front Hum Neurosci*. 2019:13:71–71. https://doi.org/10.3389/fnhum.2019.00071. https://pubmed.ncbi.nlm.nih.gov/30914935.

Tomova L, von Dawans B, Heinrichs M, Silani G, Lamm C. Is stress affecting our ability to tune into others? Evidence for gender differences in the effects of stress on self-other distinction. *Psychoneuroendocrinology*. 2014:43:95–104.

Tsakiris M. Looking for myself: current multisensory input alters self-face recognition. *PLoS One*. 2008:3:e4040.

Uddin LQ, Kaplan JT, Molnar-Szakacs I, Zaidel E, Iacoboni M. Self-face recognition activates a frontoparietal "mirror" network in the right hemisphere: an event-related fMRI study. *NeuroImage*. 2005:25:926–935.

Verbeke WJ, Rietdijk WJ, van den Berg WE, Dietvorst RC, Worm L, Bagozzi RP. The making of the Machiavellian brain: a structural MRI analysis. *J Neurosci Psychol Econ*. 2011:4:205.

Winslow JT, Insel TR. Neuroendocrine basis of social recognition. *Curr Opin Neurobiol*. 2004:14:248–253.

Wittfoth-Schardt D, Gründing J, Wittfoth M, Lanfermann H, Heinrichs M, Domes G, Buchheim A, Gundel H, Waller C. Oxytocin modulates neural reactivity to children's faces as a function of social salience. *Neuropsychopharmacology*. 2012:37:1799–1807.

Wu H, Yang S, Sun S, Liu C, Luo Y-J. The male advantage in child facial resemblance detection: behavioral and ERP evidence. *Soc Neurosci*. 2013:8:555–567.

Wu H, Feng C, Lu X, Liu X, Liu Q. Oxytocin effects on the resting-state mentalizing brain network. *Brain Imag Behav*. 2020a:14:2530–2541.

Wu H, Liu X, Hagan CC, Mobbs D. Mentalizing during social interaction: a four component model. *Cortex*. 2020b:126:242–252.

Yue T, Xu Y, Xue L, Huang X. Oxytocin weakens self-other distinction in males during empathic responses to sadness: an event-related potentials study. *PeerJ*. 2020:8:e10384.

Zhao W, Yao S, Li Q, Geng Y, Ma X, Luo L, Xu L, Kendrick KM. Oxytocin blurs the self-other distinction during trait judgments and reduces medial prefrontal cortex responses. *Hum Brain Mapp*. 2016:37:2512–2527.

Zhao W, Luo R, Sindermann C, Li J, Wei Z, Zhang Y, Liu C, Le J, Quintana DS, Montag C et al. Oxytocin modulation of self-referential processing is partly replicable and sensitive to oxytocin receptor genotype. *Prog Neuro-Psychopharmacol Biol Psychiatry*. 2020:96:109734.

Zheng S, Liang Z, Qu Y, Wu Q, Wu H, Liu Q. Kuramoto model based analysis reveals oxytocin effects on brain network dynamics. *International Journal of Neural Systems*. 2022;32:2250002.

Zheng S, Punia D, Wu H, Liu Q. Graph theoretic analysis reveals intranasal oxytocin induced network changes over frontal regions. *Neuroscience*. 2021b:459:153–165.

Zhu R, Liu C, Li T, Xu Z, Fung B, Feng C, Wu H, Luo Y, Wang L. Intranasal oxytocin reduces reactive aggression in men but not in women: a computational approach. *Psychoneuroendocrinology*. 2019:108:172–181.