SUPPLEMENTARY INFORMATION

Oxytocin reduces a chemosensory-induced stress bias in social perception

Ayline Maier, Dirk Scheele, Franny B Spengler, Tugba Menba, Franziska Mohr, Onur Güntürkün, Birgit Stoffel-Wagner, Thomas M. Kinfe, Wolfgang Maier, Sahib S. Khalsa, and René Hurlemann

SUPPLEMENTAL METHODS

Ethics and enrollment

The study was approved by the local ethics committee of the Medical Faculty of the University of Bonn, Germany. The study was registered in the Clinical Trials.gov database (Identifier: NCT03265899) provided by the US National Institutes of Health. All participants gave written informed consent and the study was conducted in accordance with the latest revision of the Helsinki Declaration. Participants were recruited from the local population by means of online advertisement and public postings. After completion of the study, participants received monetary compensation. The random allocation sequence (for the double-blind, cross-over oxytocin/placebo treatment) was generated by D.S.. A.M. enrolled all participants and assigned participants to the treatment based on the random allocation plan. All behavioral and fMRI data were collected in Bonn, Germany.

Screening session

Study enrollment was preceded by a screening session to ensure subjects were free of any current or past physical or psychiatric illness as assessed by medical history and the Mini-International Neuropsychiatric Interview (MINI) (Sheehan *et al*, 1998). Furthermore, participants were lifetime naive to prescribed psychoactive medication. None of the participants had anosmia, as verified by the Sniffin'Sticks test battery, which comprises an odor identification and discrimination test battery (Burghart GmbH, Burghart Wedel, Germany). Contraindications for MRI scanning were additional exclusion criteria. To further characterize the sample, we acquired sociodemographic data of each participant. Subjective anxiety was measured with the German version (Spielberger *et al*, 1970) of the Spielberger Trait Anxiety Inventory (STAI) and autistic-like traits were measured via the Autism Spectrum Quotient questionnaire (AQ) (Baron-Cohen *et al*, 2001) (cf. **Supplemental Table S1**).

Experimental design and procedures

We used a randomized, double-blind, PLC-controlled, within-group design. Participants self-administered 40 IU of synthetic oxytocin (OXT) or placebo (PLC) via nasal spray at the beginning of each testing session under the supervision of an experimenter and in accordance with the latest standardization guidelines (Guastella *et al*, 2013) (5 puffs balanced across nostrils, at an inter-puff interval of 50 seconds to allow the solution to be absorbed into the nasal epithelium). The amount of administered substance was weighed and was supplemented by an additional puff if it fell below a set minimum (40 IU = 1000mg). A recent study used arterial spin labeling to measure OXT-induced changes in resting regional cerebral blood flow and found that changes are sustained over a posttreatment observation interval between 25–78 min after nasal spray administration (Paloyelis et al., 2016). In a recent kinetic study, we observed decreases in amygdala responses to fearful faces between 15 and 100 min after nasal spray administration, with the strongest reduction being evident after 45 min (Spengler et al., 2017). According to this pharmacodynamic profile and given a task duration of about 20 min, we started the fMRI measurement 30 min after nasal spray administration.

For the experimental session, participants were asked to maintain their regular sleep and waking times and to abstain from caffeine and alcohol intake for 24 hours prior to study arrival. This was verified via an informational questionnaire administered at the beginning of each testing session. All female participants were required to undergo a urine pregnancy test prior to nasal spray administration. Since hormonal contraceptives have previously been shown to interfere with olfactory performance (Derntl *et al*, 2013), female participants were ask to confirm their non-use of hormonal contraception at the beginning of each testing day.

To assess possible associations between baseline OXT levels and the processing of chemosensory signals and to validate elevated peripheral OXT levels following OXT manipulation, saliva samples were collected from each subject before the intranasal administration (pre) and immediately after the MRI scan session (post), respectively. There is preliminary evidence that the fluctuations of gonadal hormones over the course of the menstrual cycle may interact with the OXT system (Kanat *et al*, 2014). Moreover, previous research has similarly shown that the olfactory sensitivity varies in different phases of the menstrual cycle and, thus, affecting the olfactory performance(Derntl *et al*, 2013). Thus, to control for

potential interaction effects, female participants were tested in their luteal phase. This was assessed by self-report and validated by blood assays (FSH, LSH, estradiol, progesterone and testosterone concentrations) obtained on the day of fMRI scanning (cf. **Supplemental Table S2**).

To evaluate for potential effects of OXT on state anxiety and mood, all subjects completed the Spielberger State Anxiety Inventory (STAI) (Spielberger *et al*, 1970) and the Positive and Negative Affective Scale (PANAS) (Watson *et al*, 1988) immediately before the OXT/ PLC administration and after the experimental task (cf. **Supplemental Table S3**). None of the participants reported olfactory problems or nasal congestion on any of the MRI testing days. MRI scanning began with the functional scan followed by an anatomical scan. The total time of an experimental session was 2h, with each participant approximately 40 min in the scanner. Participants were naïve to the purpose of the study and to the nature of the odorants presented during the fMRI task. At the end of the second experimental session, all participants were fully debriefed.

Olfactory stimuli

Sweat donors

Male axillary sweat samples were collected from donors that underwent both a bicycle ergometer training (sport condition) and the Trier Social Stress Test (TSST) (Kirschbaum *et al*, 1993) (stress condition) in two separate testing sessions a few days apart. The order of conditions was fixed across the pre-study, with all donors participating in the sport condition first. All donors were non-smokers and reported no history of somatic or mental disorders, use of medication, and drug or alcohol abuse, as assessed by the Mini-International Neuropsychiatric Interview (M.I.N.I.) (Sheehan *et al*, 1998). This pre-study was carried out in accordance with the Declaration of Helsinki principles and approved by the institutional ethics committee. Written informed consent was obtained from all donors. All donors were naïve to the applied experimental procedures. In addition, donors agreed to follow a strict protocol to prevent body odor contamination 24 h prior to the sweat donations, which included refraining from odorous food (garlic, asparagus, heavily-flavored food), alcohol, caffeine, excessive exercise, and perfumed toiletries

(perfumes, deodorants, antiperspirants, aftershaves, body lotions, shower gels, and hair products). On the sweat sampling days, donors were instructed to take a shower with a scent-free shower gel (Sensiva Washing Lotion; Schülke & Mayr GmbH, Norderstedt, Germany) that was provided by the experimenter two days prior to the first testing session and to only wear loose and odorless clothes.

Endocrine and physiological assessment

We measured the donors' free salivary cortisol levels using commercial sampling devices (Salivettes; Sarstedt AG & Co, Nümbrecht, Germany) at three time points: 10 min prior to the stressor (t1), immediately after the cessation of the stressor (t2) and 10 min after the stressor was terminated (t3). Accumulating evidence suggests that elevated free salivary cortisol levels represent an indicator of hypothalamus-pituitary-adrenal axis (HPA) activity and thus serve as a biological marker for psychosocial stress (Dickerson and Kemeny, 2004; Foley and Kirschbaum, 2010; Hellhammer *et al*, 2009). Pulse and blood pressure were measured as markers for arousal using an automatic hemodynamometer (Bosch + Sohn GmbH, Jungingen, Germany) at t1, t2, and t3. Further, sweat donors completed the STAI (Spielberger *et al*, 1970) and the PANAS (Watson *et al*, 1988) at t1 and at t2 in order to control for changes in state anxiety and mood. In addition, donors rated the level of stress they had experienced during the experimental task at t2.

Sweat donation procedures

Sweat donation sessions were carried out on an individual basis in a room in which the temperature was set at 25 °C. The sport sweat donation session consisted of 25 min of ergometer training. To maintain physiological arousal comparable between the sport stress conditions, donors were instructed to exercise at a constant pulse rate of 130 BPM, which was monitored via pulse oximetry displayed on a portable patient monitor (MEC-1200, Mindray) with the sensor clipped to the donors' left index finger. Psychosocial stress was induced utilizing the standardized protocol of the TSST. The TSST is a laboratory stressor that requires participants to prepare and deliver a speech and perform a surprise serial subtraction task in the

presence of a socially evaluative, non-responsive panel in the context of a fictitious job interview. Previous studies have demonstrated that this standardized protocol reliably induces activation of the HPA axis and thus elicits endocrine and autonomic stress responses to the perceived socio-evaluative threat (Dickerson *et al*, 2004; von Dawans *et al*, 2011). Our experimental protocol of the anticipatory speech preparation, speech delivery, and verbal arithmetic performance closely followed the standardized TSST procedures (Kirschbaum *et al.,* 1993). The panel was comprised of the same evaluators (two men, one woman) for all donors. Additionally, after the arithmetic task, donors were required to answer prepared questions with respect to analytical thinking from the panel for 5 min.

Stimuli processing

Odor stimuli were gathered from donors' axillae by attaching clean 6.5 x 4.3 x 4.3 inches cotton nursing pads (NUK Ultra Thin Disposable Nursing Pads; MAPA GmbH, Zeven, Germany) using surgical tape (Leukosilk®; BSN Medical GmbH, Luxembourg). The total period of sweat collection in both sessions was 25 min. During the donation sessions, subjects wore tight cotton t-shirts provided by the experimenter that were washed with a scent-free detergent. To avoid any bacterial contamination of the sweat pads, the experimenter wore odorless nitrile gloves at all times while handling the samples.

Sweat pads were removed immediately after stressor termination in both sessions and cut into eight pieces of the same size and put into sterile plastic bags. The pads were pooled and stored at -80°C until the olfactory stimuli were compiled for the main study. Olfactory stimuli were generated 60 min prior to the emotion recognition experiment by compiling four sweat pad pieces from four different donors in the same condition (sport or stress) in order to control for individual body scent of the donors. Non-social raspberry stimuli were prepared by pipetting 50 µl of the chemically synthesized raspberry on a quarter of a clean cotton nursing pad. For fMRI data acquisition, olfactory stimuli were filled into the glass chambers of the olfactometer and were used for two sessions. Olfactory stimuli were again stored at -80°C immediately after the fMRI sessions. In the complementary study, olfactory stimuli were defrosted 60 min prior to each experimental session and immediately frozen afterwards. Previous research demonstrated that several cycles of freezing and thawing of sweat samples do not affect the sample quality (Lenochova *et al*, 2009).

Odor delivery

The olfactometer (OG001, Burghart GmbH, Wedel, Germany) permits a precise delivery due to discrete pulses of odor with a rapid on-off time. This system is free of additional tactile, thermal or auditory cues that might interfere with the experimental task. Odorant flows (5 lpm) were directed via 10 m PTFE tubes through an odorless oxygen mask, which participants wore inside the scanner. To ensure participants received the full concentration of the odorants, we did not add any dilution air (0 lpm). Each set of olfactory stimuli (sport sweat, stress sweat, and raspberry) were used for two successive fMRI sessions. At stimuli offset, subjects breathed normal room air through the exhalation ports of the oxygen mask. Triggering of the odor channels was accomplished using a specialized olfactometer control software supplied by the manufacturer (OG Control, Burghart GmbH, Wedel, Germany).

Respiratory monitoring

Respiratory compliance was monitored on-line throughout the fMRI paradigm using a Biopac MP150 system and the accompanying AcqKnowledge Acquisition & Analysis Software (Version 4.3.1) with an MR-compatible breathing belt (RX-TSD221-MRI) affixed to the subject's chest to record thoracic contraction and expansion. The breathing belt was connected to a differential pressure transducer in the monitoring room via a 1.5 mm MR-compatible tubing (AFT30-XL) with a length of 10 m. Respiration signal was recorded applying a sampling frequency of 1000 Hz. Noise was removed by the means of hardware-based filter included in the amplifier with a low pass filter of 1 Hz and a high pass filter of 0.05 Hz. Online visual inspections of the breathing cycles revealed that all participants mastered correct breathing in more than 97% of all trials. Therefore, no data had to be excluded.

Olfactory threshold

To verify that there were no differences in olfactory functioning due to the intranasal OXT and PLC administration, a staircase olfactory threshold test (Burghart GmbH, Wedel, Germany) (Hummel *et al*, 1997) was conducted with all participants immediately after the nasal spray administrations in both testing

sessions (cf. **Supplemental Fig. S2**). Olfactory thresholds were obtained for phenylethylalcohol at low concentrations. The odor was presented with felt tip pens in 16 successive 1:2 dilution steps starting from a 4% solution. Using a three-alternative forced-choice task, three pens were presented in a randomized order at each dilution step, with two containing an odorless solvent and the third the target odorant. Subjects were instructed to identify the odor-containing pen. This triplets were presented at intervals of approximately 30 s. To prevent visual identification of the odorant-containing pens, subjects were blindfolded during the tests. Two successive correct identifications or one incorrect resulted in a reversal of the staircase. Olfactory thresholds were estimated using the mean of the last four of a total of seven staircase reversals. Only participants with normosmic scores (> 6) were included in the analyses (Hummel et al., 2007).

Olfactory stimuli ratings

Post MRI

At the end of each experimental fMRI session, participants rated the three odors they had been exposed to during the fMRI paradigm (sport sweat, stress sweat, raspberry). For this purpose, the experimenter sequentially placed the glass chambers of the olfactometer, each containing one of the stimuli, under the nostrils of the subjects. Participants were blindfolded during the odor stimuli presentation to avoid possible visual confounds. Participants evaluated the olfactory stimuli with respect to their pleasantness, intensity, and familiarity on 7-point Likert scales ranging from zero (“not at all”) to six (“very”).

Complementary study

To validate that neither social chemosensory stimuli types exhibited detectable differences in odor quality across all scan sessions, an unrelated sample of 29 healthy subjects (18 females, mean ± SD age, 24.93 ± 3.62 years) rated the pleasantness, intensity, and familiarity of the sweat stimuli after the completion of the main study. All subjects were free of physical or mental disorders as assessed by M.I.N.I and medical history. Subjects were unaware of the purpose of the study and the nature of the odors. Odor ratings were

provided on 7-point Likert scales ranging from one (“not at all”) to seven (“very”) in two separate sessions. The order of stimuli type and the order of stimuli presentation within both rating sessions were randomized across all participants. Ratings were performed on a 21.5 inches LCD monitor positioned 50 centimeters in front of the subjects, with the questions centrally aligned in black text on a white backdrop.

For each stimulus delivery, a black fixation cross was displayed during the interstimulus interval (ISI) for 1000 ms. Subjects were then cued to exhale by a red fixation cross appearing for 2000 ms on the screen. Subsequently, a green fixation cross appeared on the screen for the duration of 1500 ms, prompting the subjects to inhale. The experimenter presented the subjects with the olfactory stimuli in sterile plastic bags placed under the participants’ nostrils during the inhalation phase. Subsequently, subjects rated the olfactory stimuli by pressing the respective key number ranging from 1-7 on a keyboard positioned in front of the monitor. Subjects completed the three rating questions in a self-paced mode. Each experimental session comprised 43 trials in total and each olfactory stimulus was evaluated by 15 subjects.

Furthermore, state anxiety and salivary cortisol levels were measured at baseline and 10 min after the experiment in each rating session. Subjects were fully debriefed at the end of the second rating sessions.

fMRI paradigm

Stimuli were presented on a 32-inch MRI compatible TFT LCD monitor (NordicNeuroLab, Bergen, Norway) placed at the rear of the magnet bore using Presentation 14 (Neurobehavioral Systems, Albany, CA). Face stimuli and forced-choice questions were presented on a white backdrop. Participants were instructed to identify the emotion depicted by the face as quickly and accurately as possible by pressing one of two buttons on an MRI-compatible response grip system (NordicNeuroLab AS**,** Bergen, Norway). Participants learned the button press coding before the scan.

Visual stimuli were obtained from the stimulus set of Karolinska Directed Emotional Faces (KDEF) database (Lundqvist *et al*, 1988). To create ambiguity in facial expressions, we first generated a continuum of 20 morphing levels from prototypical neutral and fearful faces of two male actors using a morphing software (FantaMorph, Abrosoft). The resulting morph images were equally distributed between

the neutral (0% = image 1) and the fearful expression (100% = picture 20) by applying 5% increments. Since previous research (Mujica-Parodi *et al*, 2009; Wudarczyk *et al*, 2016; Zhou and Chen, 2009) reported effects of chemosensory stress signals on emotionally highly ambiguous facial expressions, we selected four face images, representing a neutral (morph level 1), a low fearful (morph level 7), a medium fearful (morph level 9) and a high fearful (morph level 20) facial expression. In a previous study (Spengler *et al*, 2017), it was shown that the majority of participants rate these low fearful stimuli as neutral rather than fearful, while this pattern of results is reversed for the medium fearful faces. In order to match the sex of both sensory modality input variables, we only selected male actors, ensuring compatibility with male axillary sweat odor signals. The resulting eight images were presented 18 times each during the experiment in a randomized sequence.

fMRI data acquisition

A Siemens MAGNETOM Trio MRI system (Siemens, Erlangen, Germany) operating at 3T and equipped with a 32-channel phased-array head coil (Siemens, Erlangen, Germany) was used to acquire T2*-weighted echoplanar (EPI) images with blood-oxygen-level-dependent contrast (TR = 2500 ms, TE = 30 ms, pixel size: 2 x 2 x 3 mm, slice thickness = 3.0 mm, distance factor = 10 %, FoV = 192 mm, flip angle = 90°, 37 axial slices). High-resolution anatomical reference images were obtained on the same scanner using a T1-weighted 3D MPRAGE sequence (imaging parameters: TR = 1660 ms, TE = 2.54 ms, matrix size: 256 x 256, pixel size: 0.8 x 0.8 x 0.8 mm, slice thickness = 0.8 mm, FoV = 256 mm, flip angle = 9°, 208 sagittal slices).

fMRI data analysis

Preprocessing

The MRI data were preprocessed and analyzed using SPM12 software (Wellcome Trust Centre for Neuroimaging, London, United Kingdom; http://www.fil.ion.ucl.ac.uk/spm) implemented in MATLAB R2010b (MathWorks, Natick, Massachusetts). The first five volumes of each functional time series were

discarded to allow for T1 equilibration. Images were corrected for head movement between scans by an affine registration. For realignment, a two pass procedure was used by which images were initially realigned to the first image of the time series and subsequently re-realigned to the mean of all images. For normalization, a two-step procedure was applied. Normalization parameters were first determined by segmenting the T1-image using the default tissue probability maps. Next, normalization parameters were applied to normalize the functional images to the standard anatomical Montreal Neurological Institute (MNI) space resampled at 2 × 2 × 2 mm voxel. The normalized images were spatially smoothed using a 6-mm FWHM Gaussian kernel. Raw time series were detrended using a high-pass filter (cut-off period, 128 s).

First-Level analysis

On the first level, 24 ('Sport_Neutral$_{OXT}$', 'Stress_Neutral$_{OXT}$', 'Raspberry_Neutral$_{OXT}$', 'Sport_LowFearful$_{OXT}$', 'Stress_LowFearful$_{OXT}$', 'Raspberry_LowFearful$_{OXT}$', 'Sport_MediumFearful$_{OXT}$', 'Stress_MediumFearful$_{OXT}$' 'Raspberry_MediumFearful$_{OXT}$', 'Sport_HighFearful$_{OXT}$', 'Stress_HighFearful$_{OXT}$', 'Raspberry_HighFearful$_{OXT}$', 'Sport_Neutral$_{PLC}$', 'Stress_Neutral$_{PLC}$', 'Raspberry_Neutral$_{PLC}$', 'Sport_LowFearful$_{PLC}$', 'Stress_LowFearful$_{PLC}$', 'Raspberry_LowFearful$_{PLC}$', 'Sport_MediumFearful$_{PLC}$', 'Stress_MediumFearful$_{PLC}$' 'Raspberry_MediumFearful$_{PLC}$', 'Sport_HighFearful$_{PLC}$', 'Stress_HighFearful$_{PLC}$', 'Raspberry_HighFearful$_{PLC}$') regressors were modeled by a stick function convolved with a hemodynamic response function, with the trial onset defined as the onset of odor delivery. The six movement regressors (realignment parameters) were included as confounds in the design matrix. A two-level random effects approach based on the general linear model as implemented in SPM12 was used for statistical analyses.

Second-Level analysis

On the second level, effects of OXT were analyzed by employing a 2 x 2 flexible factorial design with treatment (OXT, PLC) and odor (sport, stress) as within-subject factors and the BOLD-response of the contrasts [(Stress$_{(OXT)}$ – Sport$_{(OXT)}$) – (Stress$_{(PLC)}$ > Sport$_{(PLC)}$)] and [(Stress$_{(PLC)}$ > Sport$_{(PLC)}$) – (Stress$_{(OXT)}$ >

$Sport_{(OXT)}$)] as dependent variables. An additional analysis did not reveal treatment and odor interaction effects for specific emotion levels. Thus, for our main fMRI analysis, the within-subject factors were averaged across all emotion levels. In addition, to test if the task successfully resulted in stress-specific activations under PLC (i.e. the task effect), we specified the contrasts [($Stress_{(PLC)}$ > $Sport_{(PLC)}$)] and [($Sport_{(PLC)}$ > $Stress_{(PLC)}$)]. Unspecific, domain-general effects of OXT (i.e. the main effect of treatment) were analyzed by comparing all conditions with the low level baseline ([OXT>PLC]) and ([PLC>OXT]). Sex-differential OXT effects on stress-specific activations were examined using SPM independent *t*-tests.

The modulatory effect of OXT on the neural processing of the non-social control odor raspberry was analyzed with a flexible factorial design with the within-subject factor treatment (OXT, PLC) and the BOLD-response to the contrasts [$Raspberry_{(PLC)}$ > $Raspberry_{(OXT)}$] and [$Raspberry_{(OXT)}$ > $Raspberry_{(PLC)}$] as dependent variables.

To investigate whether OXT exerts a valence-specific effect on the neural processing of emotional faces, a 2 × 2 flexible factorial design with the within-subject factors treatment (OXT, PLC) and emotion level (high fearful, neutral) was performed. For this analysis, the BOLD-response of the contrasts [(high $fearful_{(OXT)}$ – $neutral_{(OXT)}$) – (high $fearful_{(PLC)}$ > $neutral_{(PLC)}$)], [(high $fearful_{(PLC)}$ > $neutral_{(PLC)}$) – (high $fearful_{(OXT)}$ > $neutral_{(OXT)}$)] were used as dependent variables. Additionally, we explored emotion-specific activations under PLC, employing the contrasts [(high $fearful_{(PLC)}$ > $neutral_{(PLC)}$)] and [($neutral_{(OXT)}$ > high $fearful_{(OXT)}$] as dependent variables.

To reflect our hypotheses derived from previous work, analyses of treatment and task effects were focused on brain structures that feature common anatomical pathways in odor and emotion processing (Gottfried, 2010; Soudry *et al*, 2011) and have been shown to exhibit responses to social olfactory stress cues in prior studies. Thus, the second level analysis focused on a set of a priori bilateral regions of interest (ROIs) including the amygdala (Mujica-Parodi *et al*, 2009), anterior cingulate cortex (Prehn-Kristensen *et al*, 2009), and hippocampus (Wudarczyk *et al*, 2016). We also selected the fusiform face area (FFA) (Kanwisher *et al*, 1997) as ROI to examine the modulatory interaction effect of OXT and chemosensory stress signals on face perception. The FFA has been demonstrated to exhibit increased activation during emotional face processing under the exposure to social chemosensory stress cues (Prehn-Kristensen *et al*, 2009; Wudarczyk *et al*, 2016). ROIs were anatomically defined according to the

Wake Forest University Pick Atlas (Version 3.0). In addition, an exploratory whole-brain analysis was performed applying a height threshold of $P < 0.001$. *P* values were corrected for multiple comparisons (family-wise error (FWE)) based on the size of the ROI and $P < 0.05$ was considered significant. To disentangle the specificity of OXT effects on social olfactory stimuli, parameter estimates were extracted from 4-mm spheres centered at the maximum *t*-value of significant clusters of the BOLD level analysis.

Connectivity analysis

A generalized psychophysiological interaction (gPPI; http://www.nitrc.org/projects/gppi) in SPM8 was used to examine the effect of OXT on functional connectivity between regions showing stress-specific activations and the FFA (Kawasaki *et al*, 2012; LaBar *et al*, 2003; Vuilleumier and Pourtois, 2007), a region that is specifically involved in encoding emotional information from faces. Compared with standard PPI implementation in SPM, gPPI methods allow for a more efficient investigation of task-dependent connectivity between identified seed regions and chosen ROIs when there are more than two task conditions (McLaren *et al*, 2012). Seed regions were identified as 4-mm radius spheres centered at the peak voxel of significant clusters for the contrasts [($Stress_{(PLC)} > Sport_{(PLC)}$)] and [($Stress_{(PLC)} > Sport_{(PLC)}$) – ($Stress_{(OXT)} > Sport_{(OXT)}$)] in the BOLD level analysis that fell within our a priori ROIs (bilateral amygdala, ACC, and hippocampus). Structurally defined ROIs of the bilateral FFA were obtained using the WFU Pickatlas. On the first level, mean time series for each condition were extracted from these 4-mm spheres and deconvolved with the hemodynamic response function (HRF). The resulting time series were multiplied with the task condition regressors and reconvolved with the HRF to obtain the PPI interaction variables. For this purpose, the same task regressors specified for the BOLD level analysis were modeled in the first level models. Separate PPI models were estimated for each seed for each participant. On the second level, obtained contrast images were entered in a 2 × 2 flexible factorial design with the within-subject factors treatment (OXT, PLC) and odor (stress, sport) averaged across emotion levels. We examined the modulatory effect of OXT on connectivity between seeds and ROIs using planned SPM dependent *t*-tests for the contrasts [($Stress_{(PLC)} > Sport_{(PLC)}$) – ($Stress_{(OXT)} > Sport_{(OXT)}$)] and [($Stress_{(OXT)} > Sport_{(OXT)}$) – ($Stress_{(PLC)} > Sport_{(PLC)}$)] in the flexible factorial design. Moreover, stress-specific functional connectivity under PLC was investigated using SPM dependent t-tests for the contrasts [($Stress_{(PLC)} >$

$Sport_{(PLC)}$)] and [($Sport_{(PLC)}$ > $Stress_{(PLC)}$)]. In a second step, we specified the contrasts [($Raspberry_{(PLC)}$ > $Raspberry_{(OXT)}$)] and [($Raspberry_{(OXT)}$ > $Raspberry_{(PLC)}$)] to investigate a potential modulatory effect of OXT on the processing of the non-social control odor raspberry.

Results were considered significant at $P_{FWE} < 0.05$ (peak-level inference) adjusted to the size of the ROIs. Parameter estimates were extracted from the 4-mm spheres centered at the maximum *t*-value of within-group differences, indicating condition-specific functional connectivity of the seed region to a target region. Connectivity results were visualized with the BrainNet Viewer (Xia *et al*., 2013).

Respiratory analysis

In order to control for potential respiratory confounds of chemosensory-induced differential neural response effects, we performed an analysis of subjects' respiratory waveforms that were acquired at the time of scanning. Respiration data were preprocessed by applying a Finite Impulse Response (FIR) Lowpass Filter with a cutoff frequency of 1 Hz. The subject-specific waveforms were baseline-adjusted with the lowest point of the entire recording curve set to zero. For each trial, the maximum and minimum waveform value and the area under the curve (AUC) were calculated, sorted by trial type and averaged in MATLAB for the subsequent statistical analyses. Maximum and minimum waveform values served as indicators for respiration depth, whereas the AUC functioned as an indicator for respiration volume. Analogous to the fMRI analysis, each trial was defined as the duration of the odor presentation and the trial onset was defined as the onset of the visual inhalation cue (green cross).

In addition, we repeated the fMRI analysis by performing a model-based respiratory noise correction. Using the MATLAB PhysIO toolbox (Kasper et al, 2017), we computed RETROICOR (retrospective image correction) regressors (Glover et al, 2000) using a 4th order Fourier expansion for the respiratory phase (Harvey et al, 2008) and RVT (respiratory volume per time; Birn et al, 2008) regressors. Noise correction was performed by entering the resulting nine nuisance regressors into the general linear model of the fMRI data.

Statistical analysis

Behavioral, demographic, and psychometric data were analyzed using SPSS 22 Version 22.0 (IBM Corp., Armonk, NY, USA). Potential demographical and neuropsychological a priori differences between both sexes were explored using independent *t*-tests. To investigate the effects of chemosensory stress signals in the emotion recognition paradigm, the mean fearful ratings and the mean response times (RTs) were analyzed in separate 2 (female, male) x 2 (OXT, PLC) x 2 (sport, stress) x 4 (neutral, low fearful, medium fearful, high fearful) mixed design analyses of variance (ANOVAs). The modulatory effect of OXT on behavioral responses under the exposure to the non-social control odor raspberry was examined in a separate 2 (female, male) x 2 (OXT, PLC) x 4 (neutral, low fearful, medium fearful, high fearful) ANOVA. Post hoc analyses to delineate higher order effects were calculated using dependent *t*-tests.

Potential differences in respiration parameters between the three odor conditions were analyzed using three separate one-way repeated measures ANOVAs for both treatment sessions.

The sweat donors' saliva cortisol levels, pulse, systolic, and diastolic blood pressure were analyzed using 2 (sport condition, stress condition) x 3 (t1, t2, t3) (rm) ANOVAs. Dependent *t*-tests for single time point evaluations were carried out post hoc. Changes in sweat donors' anxiety and subjective stress levels in the sport and stress conditions were evaluated using dependent *t*-tests.

To assess a possible treatment effect on the ratings of the olfactory stimuli, the ratings of sweat stimuli for pleasantness, intensity, and valence were evaluated in separate 2 (OXT, PLC) x 2 (sport sweat, stress sweat) rmANOVAs. In order to examine possible differences in odor quality across the two fMRI sessions irrespective of treatment, pleasantness, intensity, and valence ratings of sweat samples were compared using a 2 (sport odor stimuli in the first testing session, sport odor stimuli presentation in the second testing session) x 2 (female, male) x 2 (sport condition, stress condition) mixed-design ANOVA.

Potential changes in state anxiety and cortisol levels in participants of the complementary study in response to the chemosensory stress signals were explored by means of two separate 2 (female, male) x 2 (sport, stress) x 2 (baseline, post) mixed-design ANOVAs.

Lastly, we used Pearson's product-moment to assess whether endogenous baseline OXT levels were associated with task performance in both treatment sessions and the OXT effects on a behavioral and

neural level. We also tested whether body mass indexes predicted OXT effects on a behavioral and neural level.

The assumption of sphericity was assessed with Mauchly's test, and for significant violations Greenhouse-Geisser's correction was applied. All reported *P*-values are two-tailed and *P*-values of $P < 0.05$ were considered statistically significant. Measures of effect sizes are reported using partial eta squared values (η_p^2) for principal effects and Cohen's d (1988) for dependent and independent *t*-tests.

Brain behavior associations

To further explore associations between the neural indices of social chemosensory signal processing and emotion recognition ratings and RT, parameter estimates were extracted from 4-mm spheres centered at the maximum *t*-value of significant clusters from the BOLD level and the connectivity analysis using MarsBaR. Associations were examined using Pearson's product-moment correlation and considered significant at $P < .05$.

Hormonal assessment

Saliva samples were collected immediately before administration of the nasal spray and after the MRI scan session using pre-chilled Salivettes (Sarstedt, Rommelsdorf, Germany). Salivettes were immediately centrifuged at 3000 rpm for 2 min and aliquoted samples were stored at -80°C until assayed. Saliva OXT was extracted and quantified using a highly sensitive and specific radioimmunoassay (RIAgnosis, Munich, Germany). The limit of detection was 0.1 - 0.5 pg, depending on the age of the tracer. Intra-assay and inter-assay coefficients of variability were < 10%. All samples to be compared were assayed in the same batch, i.e. under intra-assay conditions.

In order to validate the cycle phase and to control for baseline differences in gonadal hormones levels, blood samples were collected before nasal spray administration. Serum FSH, LH and estradiol were analyzed by fully automated homogeneous sandwich chemiluminescent immunoassays based on the LOCI™ technology on a Dimension Vista™ System according to the manufacturer´s instructions (Siemens

Healthcare Diagnostics, Eschborn, Germany). The detection limits of each assay were 0.2 IU/l for LH and FSH and 11 pg/ml for estradiol. The coefficients of variation for intra-assay and inter-assay precision were <1.8 % and <2.1 % for LH, <1.9 % and <2.2 % for FSH and <5.5 % and <5.9 % for estradiol. Serum progesterone was determined by a fully automated solid-phase competitive chemiluminescent enzyme immunoassay on an Immulite™ 2000xpi System according to the manufacturer´s instructions (Siemens Healthcare Diagnostics). The detection limit of the assay was 0.1 ng/ml. The coefficients of variation for intra-assay and inter-assay precision were <4.2 % and <5.5 %. The cross-reactivity of all assays with other related compounds was minimal. Salivary testosterone was determined by a competitive enzyme immunoassay (ELISA) according to the manufacturer´s instructions (IBL International, Hamburg, Germany). The detection limit of the assay was 4.7 pg/ml. The coefficients of variation for intra-assay and inter-assay precision were <7.1 % and <7.7 %.

In the sweat donation prestudy, saliva samples were collected at three different time points. Immediately after the donation session, salivettes were centrifuged at 3000 rpm for 2 min and aliquoted samples were then stored at -80°C until assayed. Cortisol concentrations were determined using an electrochemiluminence immunoassay (Elecsys Cortisol Test, Roche, Mannheim).The sensitivity of the assay was set at 0.018-63.4 µg/dl. The mean inter- and intra-assay coefficient of variation for the assays were 3.42% and 12.2%, respectively.

SUPPLEMENTARY RESULTS

Missing values

RT of one female participant were not recorded in the PLC condition due to technical issues. For three female participants, the olfactory threshold could not be determined due to an erroneous test execution. Ten OXT saliva samples were lost due to problems in sample assessment or analysis, resulting in 48 (24 females) salivary OXT concentration comparisons.

Demographical and neuropsychological data, salivary OXT concentrations

Female and male participants exhibited no differences of demographical or neuropsychological measurements (all *P*s > 0.05; cf. **Supplementary Table S1**). Intranasal treatment had no effect on mood or state anxiety in both sexes (cf. **Supplemental Table S3**). Salivary OXT concentrations did not differ between treatment sessions at baseline, but were significantly elevated following OXT compared to PLC treatment (cf. **Supplemental Table S2**). None of the participants reported any side effects after treatment.

Pre-study results

Salivary cortisol levels, diastolic and systolic blood pressure, and pulse rate were analyzed using 2 × 3 repeated measures (rm)ANOVAs with the within-subject factors condition (sport, stress) and time point (pre, post, 10 min post). For endocrine responses, the analysis yielded significant main effects of condition ($F_{(1, 29)}$ = 13,73, P = 0.001, η_p^2 = 0.32) and time point ($F_{(2, 33.85)}$ = 7.11, P = 0.002 η_p^2 = 0.20) and a significant condition × time point interaction effect ($F_{(1.30, 37.58)}$ = 12.38, P < 0.001, η_p^2 = 0.30). Post hoc t tests confirmed that cortisol levels were significantly more elevated after the stress condition ($t_{(29)}$ = -4,43, P < 0.001, d = 0.8) relative to the sport condition, with this effect still being present 10 min after stressor termination ($t_{(29)}$ = 6.44, P < 0.001, d = 1.20, cf. **Fig. 1**). Pulse rate analysis revealed significant main effects of condition ($F_{(1, 29)}$ = 40.02, P < 0.001, η_p^2= 0.58) and time point ($F_{(2, 47.25)}$ = 63.48, P < 0.001, η_p^2 =

0.69) as well as a significant condition × time point interaction effect ($F_{(2, 49.25)}$ = 44.01, P < 0.001, η_p^2 = 0.60). A post hoc *t* test showed that sweat donors exhibited significantly higher pulse rates after the sport condition ($t_{(29)}$ = 8.24, P < 0.001, d = 1.50) compared to the stress condition. This effect could also be observed 10 min after the sweat donation sessions had ended ($t_{(29)}$ = 5.26, P < 0.001, d = 0.96, cf. **Supplementary Fig. S1**, reflecting the physical exertion of the sport condition. For systolic blood pressure measurements, the analysis yielded significant main effects of condition ($F_{(1, 29)}$ = 7.38, P < 0.011, η_p^2 = 0.20) and time point ($F_{(1.89, 54.80)}$ = 32.97, P < 0.001, η_p^2 = 0.53) with no interaction effect (P > 0.30). Post hoc *t* tests revealed that sweat donors exhibited significantly higher systolic blood pressure immediately after the stress condition ($t_{(29)}$ = -2.19, P = 0.04, d = -0.40) and 10 min after the stress condition ($t_{(29)}$ = -2.58, P = 0.02, d = 0.47) relative to the sport condition. The analysis of the diastolic blood pressure yielded significant main effects of condition ($F_{(1, 29)}$ = 17.55, P < 0.001, η_p^2 = 0.38) and time point ($F_{(1.90, 55.11)}$ = 28.51, P < 0.001, η_p^2 = 0.50), as well as a significant condition × time point interaction effect ($F_{(1.66, 48.08)}$ = 9.66, P = 0.001; η_p^2 = 0.25). Post hoc *t* tests confirmed that donors had significantly higher diastolic blood pressures immediately after the stress condition ($t_{(29)}$ = -6.55, P < 0.001, d = 1.20) and 10 min after the stress condition ($t_{(29)}$ = -3.09, P = 0.004, d = 0.56, cf. **Supplemental Fig. S1**) compared to the sport condition. As expected, salivary cortisol levels, blood pressure and pulse rates were comparable between both sweat donation conditions at baseline measurements (all Ps > 0.3).

Moreover, the stress condition relative to the sport condition evoked significantly higher state anxiety levels in sweat donors ($t_{(29)}$ = -8.61, P < 0.001, d = 1.57). Consistently, donors evaluated the degree of experienced stress as significantly more pronounced in the stress condition compared to the sport condition ($t_{(29)}$ = -2.32, P = 0.03, d = 0.42, cf. **Fig. 1**). Thus, stress induction utilizing the TSST resulted in a robust increase in physiological and psychological stress responses, indicating that our stress manipulation was successful.

Odor stimuli ratings

Post MRI ratings

Analyses of post MRI pleasantness, intensity, and familiarity ratings of the three chemosensory stimuli did not reveal any significant differences between treatment sessions (all *P*s > 0.05; **cf. Supplemental Fig. S2**).

Complementary study

Three separate 2 × 2 × 2 mixed-design ANOVAs yielded no main effects of sweat donation condition, sex, stimuli presentation order, or interaction effects on pleasantness ratings (all *P*s > 0.05), intensity ratings (all *P*s > 0.05), or familiarity ratings (all *P*s > 0.05, cf. **Supplemental Tables S4, S5**). These results suggest that social chemosensory stimuli used during fMRI sessions were comparable with no consciously distinguishable odor qualities across all scan sessions. This pattern of results precludes simple odor discrimination as an explanation for the stress-specific behavioral bias and neural activation.

State anxiety and cortisol levels in the complementary study

Results showed that chemosensory stress signals did not elicit changes in salivary cortisol levels in subjects of the complementary study, as the analysis yielded no main or interaction effects (all *P*s > 0.05). Concomitantly, there was a significant main effect of time point on reported state anxiety ($F_{(1, 27)}$ = 6.31, *P* = 0.02, η_p^2 = 0.19) but no interaction effect (*P* > 0.05). Thus, rating stress sweat did not evoke measureable emotional or peripheral stress reactions.

Olfactory threshold

Paired *t*-tests demonstrated that olfactory thresholds remained unchanged by nasal spray administration in both men ($t_{(27)}$ = 0.21, *P* = 0.84, *d* = 0.04) and women ($t_{(26)}$ = 1.77, *P* = 0.09, *d* = 0.37). Thus, behavioral

and fMRI findings of the present study cannot be attributed to differences in olfactory functioning as a result of intranasal spray administration. OXT did not alter the processing of the non-social odor on the behavioral or neural level (cf. SI). OXT had no effect on state anxiety and mood ratings (cf. SI).

Emotion recognition and response time

Social chemosensory signals

As a first step, to explore potential sex differences in emotion recognition ratings and RT, we performed two separate 4-way mixed-design ANOVAs. Analyses yielded a significant main effect of intensity ($F_{(1.74, 95.82)} = 373.93$, $P < 0.001$, $\eta_p^2 = 0.87$) on emotion recognition and on RTs ($F_{(2.65, 142.91)} = 40.18$, $P < 0.001$, $\eta_p^2 = 0.43$). Furthermore, analyses revealed significant sex × treatment × sweat donation condition × intensity interaction effects on emotion recognition ($F_{(3, 165)} = 4.12$, $P = 0.008$, $\eta_p^2 = 0.07$) and a similar, non-significant, interaction effect on RT ($F_{(2.53, 136.39)} = 2.02$, $P = 0.12$, $\eta_p^2 = 0.04$). In the second step, to break down the complex four-way interactions, we examined the effects of treatment, odor, and intensity in separate repeated-measures 2 × 2 × 4 repeated-measures (rm) ANOVAs at both levels of sex.

Non-social chemosensory signals

Using the emotion recognition rating as the dependent variable, a 2 × 2 × 4 mixed-design ANOVA with the between-subject factor sex and the within-subject factors treatment and intensity yielded a significant main effect of intensity ($F_{(1.83, 379.26)} = 359.31$, $P < 0.001$, $\eta_p^2 = 0.87$) and an interaction of intensity and sex ($F_{(1.80, 98.73)} = 3.88$, $P = 0.03$, $\eta_p^2 = 0.07$), but no further main or interaction effects ($P > 0.05$). Higher fear intensity was associated with a higher proportion of fear recognition, with women displaying a steeper increase than men. Using RT as the dependent variable, analysis showed a significant main effect of intensity ($F_{(2.64, 142.32)} = 29.14$, $P < 0.001$, $\eta_p^2 = 0.35$) and a significant main effect of sex ($F_{(1, 54)} = 4.26$, $P = 0.04$, $\eta_p^2 = 0.07$) and an interaction of treatment and sex ($F_{(1, 54)} = 4.46$, $P = 0.04$, $\eta_p^2 = 0.07$). Under PLC, women needed significantly less time for the emotion recognition ratings than men did ($F_{(1, 55)} = 7.84$, $P < 0.01$, $\eta_p^2 = 0.13$), while this difference vanished after OXT treatment ($P = 0.36$).

fMRI results

There was no main effect of treatment [(OXT > PLC) and (PLC > OXT)], indicating that OXT had no unspecific global effects. There were no associations between neural indices of stress sweat processing and emotion recognition ratings (all *P*s > 0.05).

Sex differences in BOLD level analysis

All participants exhibited distinct neural activation patterns in response to chemosensory stress signals. Notably, independent two-sample t-tests revealed that the stress-specific activation were not significantly different between women and men ([(Stress$_{(PLC)}$ > Sport$_{(PLC)}$)] and [(Sport$_{(PLC)}$ > Stress$_{(PLC)}$)]) for any of the pre-selected ROIs (all *Ps* > 0.05). Furthermore, there were no sex-specific treatment effects ([(Stress$_{(OXT)}$ – Sport$_{(OXT)}$) – (Stress$_{(PLC)}$ > Sport$_{(PLC)}$)], [(Stress$_{(PLC)}$ > Sport$_{(PLC)}$) – (Stress$_{(OXT)}$ > Sport$_{(OXT)}$)]) evident across all ROIs (all *P*s > 0.05).

The modulatory effect of OXT on raspberry odor processing

Whole-brain analysis showed that under PLC the combined presentation of raspberry odor and emotional faces relative to the low-level baseline elicited increased activation in a large-sized cluster extending from the right lingual gyrus (22, -82, -10; $t_{(98)}$= 16.46, k_E= 26 854, P_{FWE} < 0.001) to the bilateral calcarine sulcus (16, -92, -2; $t_{(98)}$= 16.24; 16, -92, -2; $t_{(98)}$= 15.58) across all participants. We also found increased activation in the right middle cingulum (0,-28, 26; $t_{(98)}$= 6.11, k_E= 175, P_{FWE} = 0.021). Furthermore, a ROI-based analysis revealed raspberry-induced activation in the right ACC (18, 42, 14; $t_{(98)}$= 3.73, P_{FWE} = 0.037), the bilateral hippocampus (24, -28, -6; $t_{(98)}$= 8.99, P_{FWE} < 0.001; 20, -4, -14; $t_{(98)}$= 5.57, P_{FWE} < 0.001; 22, -34, 2; $t_{(98)}$= 4.94, P_{FWE} < 0.001; -22, -30, -4; $t_{(98)}$= 10.10, P_{FWE} < 0.001; -18, -6, -12; $t_{(98)}$= 4.91, P_{FWE} = 0.001), the bilateral amygdala (22, -2, -14; $t_{(98)}$= 6.75, P_{FWE} < 0.001; -22, -4, -12; $t_{(98)}$= 6.70, P_{FWE} < 0.001), and the bilateral FFA (26, -74, -10; $t_{(98)}$= 15.31, P_{FWE} < 0.001; 35, -54, -12; $t_{(98)}$= 11.26; P_{FWE} < 0.001; -24, -72, -8; $t_{(98)}$= 13.22; -22, -80, -8; $t_{(98)}$= 13.11, P_{FWE} < 0.001; -18, -88, -8; $t_{(98)}$= 12.55, P_{FWE} < 0.001). There was no treatment effect ([(Raspberry$_{(OXT)}$ > Raspberry$_{(PLC)}$)]), ([(Raspberry$_{(PLC)}$ > Raspberry$_{(OXT)}$)]), when

participants received the non-social control odor raspberry.

Female participants

For females, whole-brain analysis revealed increased activation in large-sized cluster extending from the left calcarine sulcus (-8, -90, -6; $t_{(50)}$= 14.01, k_E= 13641, P_{FWE} < 0.001) to the right lingual gyrus (24, -82, -10; $t_{(50)}$= 13.22; 18, -90, -8; $t_{(50)}$= 13.14) and in a large-sized cluster comprising the left insula (-38, 18, -6; k_E= 17780, $t_{(50)}$= 6.18, P_{FWE} < 0.001), the right inferior frontal gyrus (48, 20, -10; $t_{(50)}$= 5.44), and the left superior temporal pole (-50, 16, -10; $t_{(50)}$= 5.43) in response to raspberry under PLC. Moreover, there were also heightened neural responses in a medium-sized cluster in the left hippocampus (-24, -24, -8; $t_{(50)}$= 8.64, k_E= 409, P_{FWE} < 0.001; -36, -24, -10; $t_{(50)}$= 3.97).

A ROI-based approach showed that raspberry odor elicited increased neural responses in the bilateral hippocampus (24, -28, -6; $t_{(50)}$= 6.12, P_{FWE} < 0.001; 16, -28, -6; $t_{(50)}$= 4.48, P_{FWE} < 0.006;-24, -24, -8; $t_{(50)}$= 8.64, P_{FWE} < 0.001; -30, -32, -2; $t_{(50)}$= 5.39, P_{FWE} < 0.001; -36, -24, -10; $t_{(50)}$= 3.97, P_{FWE} =0.025), the bilateral amygdala (22, -4, -14; $t_{(50)}$= 4.78, P_{FWE} = 0.001; 30, 4, -18; $t_{(50)}$= 3.23, P_{FWE} = 0.048), and in the bilateral FFA (26, -74, -8; $t_{(50)}$= 12.47, P_{FWE} < 0.001; 26, -82, -10; $t_{(50)}$= 11.47, P_{FWE} < 0.001; 38, -58, -14; $t_{(50)}$= 8.87, P_{FWE} < 0.001; -24, -74, -8; $t_{(50)}$= 11.80, P_{FWE} < 0.001; -20, -84, -8; $t_{(50)}$= 10.49, P_{FWE} < 0.001; -38, -42, -18; $t_{(50)}$= 9.81, P_{FWE} < 0.001) under PLC. As expected, we did not find any modulatory effects of OXT on the processing of raspberry odor in females.

Male participants

For males, whole-brain analysis exhibited increased activation in a large-sized cluster comprising the left calcarine sulcus (-6, -88, -2; $t_{(50)}$= 12.62, k_E= 19468, P_{FWE} < 0.001) to the right lingual gyrus (14, -90, -6; $t_{(50)}$= 11.87) and the left middle occipital gyrus (-16, -96, 6; $t_{(50)}$= 11.40). Further, we found elevated responses in a cluster comprising the right inferior frontal gyrus (48, 20, -10; $t_{(50)}$= 7.65, k_E= 1653, P_{FWE} < 0.001; 40, 34, -2; $t_{(50)}$= 6.25) and the right putamen (32, 20, 0; $t_{(50)}$= 6.15), and in a cluster ranging from the left insula (-36, 20, -2; $t_{(50)}$= 6.41, k_E= 911, P_{FWE} < 0.001) to the left inferior frontal gyrus (-48, 18, -10; $t_{(50)}$= 5.77) and the left putamen (-20, 20, -4; $t_{(50)}$= 5.21) under PLC in men.

ROI-analysis showed that under PLC males exhibited increased activation in the bilateral hippocampus (peak MNI coordinates x, y, z: 22, -28, -6; $t_{(50)}$= 7.39, P_{FWE} < 0.001; 20, -34, 4; $t_{(50)}$= 4.44, P_{FWE} = 0.006; 34, -18, 8; $t_{(50)}$= 4.13, P_{FWE} = 0.015; 20, -4, -12; $t_{(50)}$= 4.14, P_{FWE} = 0.015; -22, -30, -4; $t_{(50)}$= 6.62, P_{FWE} < 0.001), the bilateral amygdala (22, -2, -12; $t_{(50)}$= 4.39, P_{FWE} = 0.004; -26, 0, -16; $t_{(50)}$= 5.79, P_{FWE} < 0.001), and in the bilateral FFA (26, -72, -10; $t_{(50)}$= 9.63, P_{FWE} < 0.001; 36, -54, -12; $t_{(50)}$= 9.14, P_{FWE} < 0.001; 40, -44, -16; $t_{(50)}$= 6.50, P_{FWE} < 0.001; -22, -80, -8; $t_{(50)}$= 9.20, P_{FWE} < 0.001; -26, -72, -8; $t_{(50)}$= 8.14, P_{FWE} < 0.001; -18, -88, -8; $t_{(50)}$= 8.05, P_{FWE} < 0.001) in response to raspberry odor after PLC administration. However, the analysis did not reveal any modulatory effect of OXT on the neural processing of raspberry in males for any of the pre-selected ROIs.

As expected, raspberry odor induced strong neural responses in the olfactory network in females and males under PLC. However, the analysis demonstrated that OXT does not exert a modulatory effect on non-social chemosensory signals. These neural findings are paralleled by an absent modulatory effect of OXT on the behavioral level, when participants were exposed to raspberry odor.

The modulatory effect of OXT in response to social chemosensory signals

Amygdala and hippocampus seeds did not exhibit altered functional coupling with the FFA (all *P*s > 0.05) under stress sweat exposure. In male participants, we detected an increased functional coupling of the right ACC (12, 42, 2; 6, 48, 8) and the left FFA (-28, -60, -10; $t_{(92)}$ = 3.92; P_{FWE} = 0.047) when they were exposed to stress sweat under OXT [($Stress_{(OXT)}$ > $Sport_{(OXT)}$)].

OXT had no unspecific global effects on functional connectivity, that is there was a main effect of treatment [(OXT > PLC) and (PLC > OXT)] neither across all participants nor in separate analyses for female and male participants.

Respiratory confounding effects on fMRI data

The respiration waveform analysis revealed no difference in respiration parameters between the three odor conditions in both treatment sessions across all participants and within each sex group. Thus, there were no significant variations in respiration depth and respiration volume (all *P*s > 0.05).

Moreover, respiratory noise correction yielded only marginal and negligible changes compared to our original findings. The originally observed fMRI task effects, modulatory OXT effects and connectivity findings remained unchanged in both sex groups. Thus, is seems highly unlikely that respiration-related effects biased our fMRI findings.

Further moderation effects

For females, correlational analyses yielded a significant association of baseline OXT levels and the emotion recognition of neutral faces in the OXT condition (r = 0.50, P = 0.007). This relationship was also present using the mean of baseline OXT levels of both sessions (r = 0.50, P = 0.007). Relatedly, we observed correlations between the baseline OXT level in the OXT session (r = -0.44, P = 0.02) and the mean baseline level (r = -0.39, P = 0.04) with the OXT effect for neutral faces. However, visual inspection of the scatterplots revealed that these correlations were driven by an outlier (z > 2.5 for the task effect). After excluding one female participant from the analysis, these correlations failed to reach significance (all *P*s > 0.05). Moreover, we did not observe any further significant associations between endogenous baseline OXT levels and task performance in both treatment sessions on the behavioral and neural level in females or males.

For males, correlational analysis showed a significant relationship between body mass index (BMI) and the OXT effect on emotion recognition of high fearful faces [$(Stress_{(OXT)} – Sport_{(OXT)}) – (Stress_{(PLC)} > Sport_{(PLC)})$] ($r$ = 0.37, P = 0.049). However, visual inspection of the scatterplot revealed that this correlation was driven by an outlier (BMI: z > 2.5). After excluding one male participant from the analysis, this correlation was no longer significant (r = 0.05, P = 0.792). We did not observe any further significant associations between body mass indexes and neural and behavioral OXT effects for males and females

(all *P*s > 0.05). Thus, the body mass indexes did not moderate the neural or the behavioral OXT effect in our study.

Supplemental tables

Table S1. Demographics and psychometric trait data

	Females (n = 30) Mean (± SD)	Males (n = 28) Mean (± SD)	*t*	*P*
Age (years)	25.23 (± 3.06)	24.54 (± 3.18)	0.85	0.54
Education (years)	17.50 (± 3.11)	16.25 (± 2.39)	1.71	0.09
STAI trait [b]	33.07 (± 7.02)	29.71 (± 6.60)	1.87	0.07
AQ [c]	13.40 (± 5.64)	14.25 (± 4.21)	-0.16	0.52

Notes. Trait anxiety symptoms were assessed by the [b] STAI (State Trait Anxiety inventory). Autistic-like traits were assessed by the [c] AQ (Autism Spectrum Quotient).

Table S2. Baseline measurement of endocrine factors and olfactory thresholds

	OXT session (n = 30) Mean (± SD)	PLC session (n = 30) Mean (± SD)	*t*	*P*
Females				
Baseline Oxytocin (pg/ml)	1.56 (± 0.61)	1.41 (± 0.58)	1.10	0.28
Post MRI Oxytocin (pg/ml)	38.15 (± 28.20)	2.26 (± 0.93)	6.24	>0.001
Estradiol (pg/ml)	108.30 (± 62.81)	96.37 (± 66.73)	1.46	0.16
FSH (U/l)	4.14 (± 2.34)	4.37 (± 1.81)	-0.72	0.48
LH (U/l)	9.93 (± 14.77)	7.41 (± 6.04)	1.09	0.29
Progesterone (ng/ml)	4.37 (± 3.68)	3.90 (± 4.28)	0.73	0.48
Testosterone (pg/ml)	0.29 (± 0.14)	0.30 (± 0.18)	-0.17	0.87
Olfactory threshold	12.84 (± 2.32)	11.99 (± 2.22)	1.77	0.09
	OXT session (n = 28) Mean (± SD)	PLC session (n = 28) Mean (± SD)	*t*	*P*
Males				
Baseline Oxytocin (pg/ml)	1.10 (± 0.75)	1.00 (± 0.69)	0.42	0.68
Post MRI Oxytocin (pg/ml)	29.02 (± 23.37)	1.27 (± 0.85)	5.82	>0.001
Estradiol (pg/ml)	29.44 (± 7.12)	27.79 (± 5.53)	0.84	0.42
FSH (U/l)	3.27 (± 2.04)	3.62 (± 2.15)	-0.90	0.39
LH (U/l)	4.65 (± 2.42)	4.95 (± 2.29)	-0.82	0.43
Progesterone (ng/ml)	0.24 (± 0.13)	0.21 (± 0.10)	0.75	0.47
Testosterone (pg/ml)	4.14 (± 1.26)	3.85 (± 0.97)	1.50	0.16
Olfactory threshold	12.03 (± 2.72)	11.92 (± 3.13)	0.21	0.84

Notes. FSH, follicle-stimulating hormone; LH, luteinizing hormone; OXT, oxytocin; PLC, placebo.

Table S3. State measurements of anxiety, mood and attention

	OXT session (n = 58) Mean (± SD)	PLC session (n = 58) Mean (± SD)	t	P
STAI state pre [a]	32.14 (6.14)	33.22 (7.43)	1.15	0.26
STAI state post [a]	31.55 (6.49)	32.82 (7.1)	1.36	0.18
PANAS positive pre [b]	34.52 (5.99)	33.65 (5.61)	1.53	0.13
PANAS positive post [b]	33.59 (6.98)	33.36 (6.69)	0.44	0.67
PANAS negative pre [b]	15.52 (5.06)	14.86 (4.79)	1.32	0.19
PANAS negative post [b]	14.24 (4.75)	14.28 (4.44)	-0.06	0.95

Notes. State anxiety before and after the experiment was assessed using the [a]STAI = State Trait Anxiety Inventory. Mood before and after the experiment was assessed using the [b] PANAS = Positive and Negative Affect Schedule. Abbreviations: OXT, oxytocin; PLC, placebo.

Table S4. **Ratings of chemosensory stimuli after the fMRI sessions: women**

	PLC session Mean (± SD)	OXT session Mean (± SD)	*t*	*P*
Sport odor (Ergometer training)				
Pleasantness	2.17 (1.62)	2.40 (1.59)	-0.82	0.42
Intensity	2.97 (1.52)	3.17 (1.53)	0.61	0.55
Familiarity	2.23 (1.48)	2.70 (1.56)	1.27	0.21
Stress odor (TSST)				
Pleasantness	2.43 (1.50)	2.07 (1.72)	-0.98	0.33
Intensity	2.97 (1.43)	3.10 (1.56)	0.40	0.69
Familiarity	2.70 (1.49)	2.80 (1.52)	0.30	0.77
Non-social control odor (Raspberry)				
Pleasantness	4.23 (1.38)	4.33 (1.58)	0.43	0.67
Intensity	5.37 (0.72)	5.57 (0.57)	1.65	0.11
Familiarity	5.23 (0.97)	5.17 (1.37)	-0.21	0.84

Notes. Females' pleasantness, intensity, and familiarity ratings (scale 0-6) of the chemosensory stimuli they were exposed to during the emotion recognition fMRI paradigm under PLC (placebo) and OXT (oxytocin). Social chemosensory stimuli were male axillary sweat samples collected from an unrelated sample in a pre-study in two emotionally distinct situations: a sport condition consisting of ergometer training and a stress condition consisting of a TSST (Trier Social Stress Test). Chemically synthesized raspberry concentrate was used as a non-social control odor.

Table S5. **Ratings of chemosensory stimuli after the fMRI sessions: men**

	PLC session Mean (± SD)	OXT session Mean (± SD)	*t*	*P*
Sport odor (Ergometer training)				
Pleasantness	2.57 (1.81)	2.61 (1.45)	0.12	0.91
Intensity	2.57 (1.45)	2.82 (1.22)	0.69	0.50
Familiarity	2.96 (1.69)	2.68 (1.25)	-0.94	0.36
Stress odor (TSST)				
Pleasantness	2.12 (1.29)	2.68 (1.59)	1.77	0.09
Intensity	2.68 (1.54)	2.64 (1.31)	-0.10	0.92
Familiarity	2.79 (1.42)	2.61 (1.42)	-0.66	0.52
Non-social control odor (Raspberry)				
Pleasantness	4.43 (1.10)	4.32 (1.47)	-0.57	0.57
Intensity	5.11 (0.79)	5.21 (0.92)	0.59	0.56
Familiarity	4.75 (1.32)	4.96 (1.23)	1.36	0.18

Notes. Males' pleasantness, intensity, and familiarity ratings (scale 0-6) of the chemosensory stimuli they were exposed to during the emotion recognition fMRI paradigm under PLC (placebo) and OXT (oxytocin). Social chemosensory stimuli were male axillary sweat samples collected from an unrelated sample in a pre-study in two emotionally distinct situations: a sport condition consisting of ergometer training and a stress condition consisting of a TSST (Trier Social Stress Test). Chemically synthesized raspberry concentrate was used as a non-social control odor.

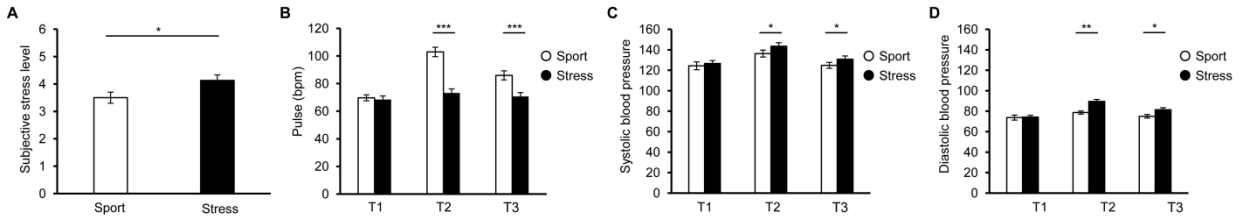
Table S6. **Evaluation of social chemosensory stimuli in the complementary study**

	Sport condition (Ergometer training) Mean (± SD)	Stress condition (TSST) Mean (± SD)	*t*	*P*
Females (n = 18)				
Pleasantness	2.92 (0.94)	2.67 (0.83)	-1.62	0.12
Intensity	3.02 (1.15)	3.14 (1.07)	0.89	0.39
Familiarity	3.45 (0.78)	3.43 (0.55)	-0.10	0.92
	Sport condition (Ergometer training) Mean (± SD)	Stress condition (TSST) Mean (± SD)	*t*	*P*
Males (n =11)				
Pleasantness	3.22 (0.91)	3.01 (0.91)	-1.76	0.11
Intensity	3.14 (0.69)	3.08 (0.94)	-0.37	0.72
Familiarity	3.24 (0.39)	3.27 (0.60)	0.15	0.87

Notes. Pleasantness, intensity, and familiarity ratings (scale 0-6) of chemosensory stimuli that were presented during the emotion recognition fMRI paradigm by an unrelated sample of 29 healthy participants (mean ± SD age, 24.93 ± 3.62 years) in a complementary study. Social chemosensory stimuli were male axillary sweat samples collected from an unrelated sample in a pre-study in two emotionally distinct situations: a sport condition consisting of ergometer training and a stress condition consisting of a TSST (Trier Social Stress Test).

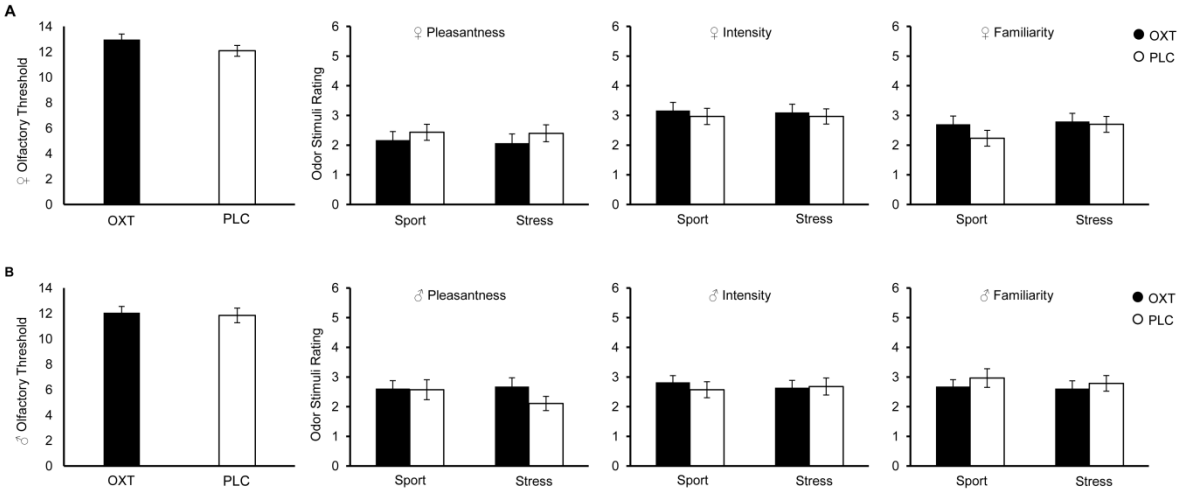
Supplemental figures

Figure S1.



Psychological and physiological measurements of the sweat donors. Donors reported higher subjective stress after the Trier Social Stress Test compared to the sport condition ($t_{(29)}$ = -2.32, P = 0.03, d = 0.42; **A**). Pulse, systolic and diastolic blood pressure measurements at baseline (T1), immediately after the stressor (T2), and 10 min after stressor termination (T3). Sweat donors exhibited higher pulse rates in the sport condition at T2 ($t_{(29)}$ = 8.24, P < 0.001, d = 1.50) and T3 ($t_{(29)}$ = 5.26, P < 0.001, d = 0,96) compared to the stress condition (**B**). Systolic (**C**) and diastolic blood pressure (**D**) measurements on the other hand were greater after the stress sweat donation session at T2 ($t_{(29)}$ = -2.19, P = 0.04, d = -0.40; $t_{(29)}$ = -6.55, P < 0.001, d = 1.20) and T3 ($t_{(29)}$ = -2.58, P = 0.02, d = 0.47; $t_{(29)}$ = -3.09, P = 0.004, d = 0.56). Error bars represent the standard error of mean (SEM). ***P < 0.001; **P < 0.01; *P < 0.05.

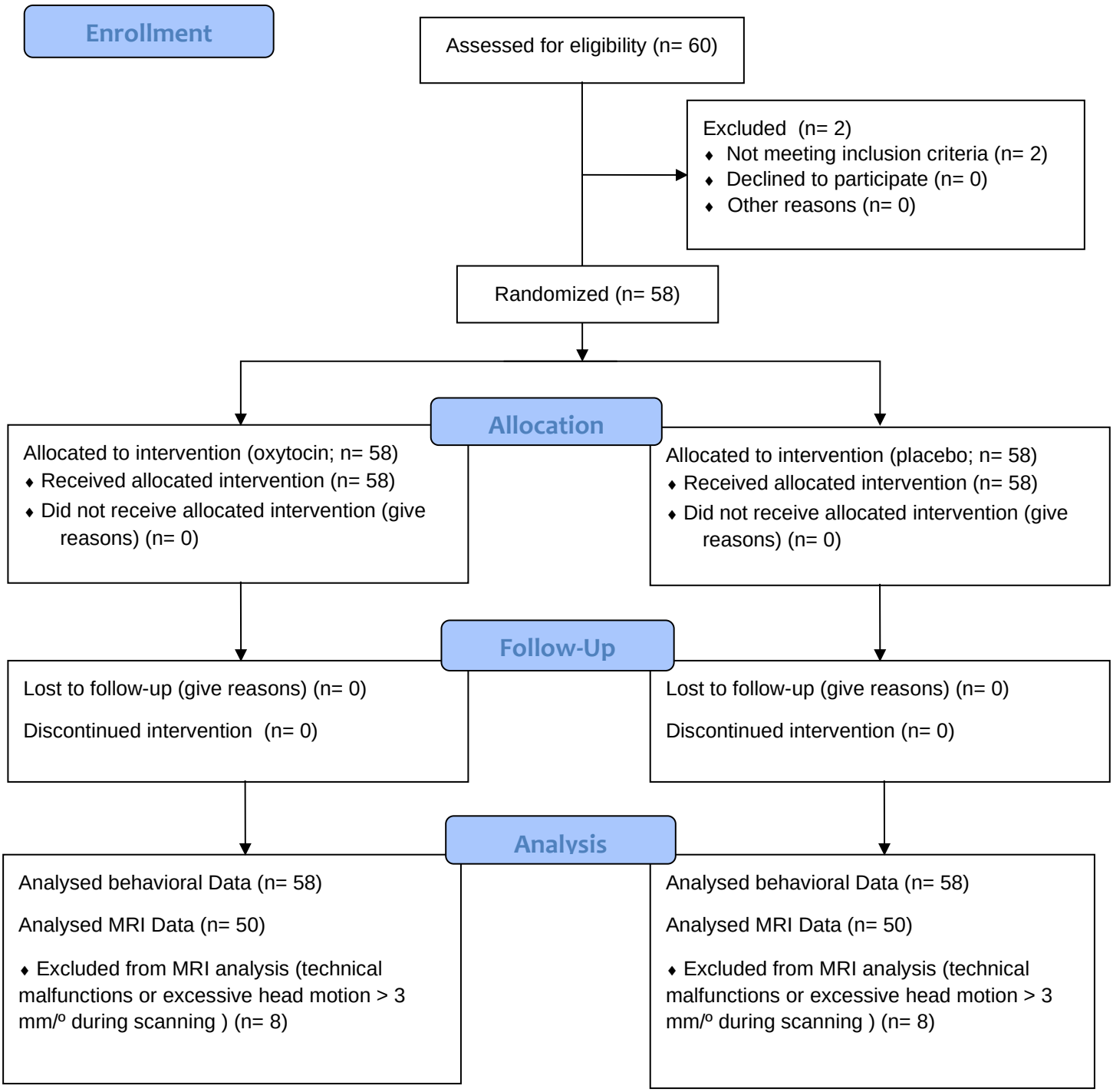
Figure S2.



Olfactory threshold and olfactory stimuli ratings in women (**A**) and men (**B**). Nasal spray treatment did not affect the olfactory threshold in female and male participants. Female and male participants rated social olfactory stimuli (sport sweat, stress sweat) similar with respect to perceived pleasantness, intensity, and familiarity in the OXT and PLC condition. Error bars represent the standard error of mean (SEM). Abbreviations: OXT, oxytocin; PLC, placebo.

CONSORT
TRANSPARENT REPORTING of TRIALS

CONSORT 2010 Flow Diagram



Supplemental references

Baron-Cohen S, Wheelwright S, Skinner R, Martin J, Clubley E (2001). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J Autism Dev Disord* **31**(1): 5-17.

Birn RM, Smith MA, Jones TB, Bandettini PA (2008). The respiration response function: the temporal dynamics of fMRI signal fluctuations related to changes in respiration. *NeuroImage* **40**(2): 644-654.

Derntl B, Schopf V, Kollndorfer K, Lanzenberger R (2013). Menstrual cycle phase and duration of oral contraception intake affect olfactory perception. *Chem Senses* **38**(1): 67-75.

Dickerson SS, Kemeny ME (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychol Bull* **130**(3): 355-391.

Foley P, Kirschbaum C (2010). Human hypothalamus-pituitary-adrenal axis responses to acute psychosocial stress in laboratory settings. *Neurosci Biobehav Rev* **35**(1): 91-96.

Glover GH, Li TQ, Ress D (2000). Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn Reson Med* **44**(1): 162-167.

Gottfried JA (2010). Central mechanisms of odour object perception. *Nat Rev Neurosci* **11**(9): 628-641.

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**(5): 612-625.

Hellhammer DH, Wust S, Kudielka BM (2009). Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology* **34**(2): 163-171.

Hummel T, Sekinger B, Wolf SR, Pauli E, Kobal G (1997). 'Sniffin' sticks': olfactory performance assessed by the combined testing of odor identification, odor discrimination and olfactory threshold. *Chem Senses* **22**(1): 39-52.

Kanat M, Heinrichs M, Domes G (2014). Oxytocin and the social brain: neural mechanisms and perspectives in human research. *Brain Res* **1580**: 160-171.

Kanwisher N, McDermott J, Chun MM (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* **17**(11): 4302-4311.

Kasper L, Bollmann S, Diaconescu AO, Hutton C, Heinzle J, Iglesias S, et al (2017). The PhysIO Toolbox for Modeling Physiological Noise in fMRI Data. J Neurosci Methods **276**: 56-72.

Kawasaki H, Tsuchiya N, Kovach CK, Nourski KV, Oya H, Howard MA*, et al* (2012). Processing of facial emotion in the human fusiform gyrus. *J Cogn Neurosci* **24**(6): 1358-1370.

Kirschbaum C, Pirke KM, Hellhammer DH (1993). The 'Trier Social Stress Test'--a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology* **28**(1-2): 76-81.

LaBar KS, Crupain MJ, Voyvodic JT, McCarthy G (2003). Dynamic perception of facial affect and identity in the human brain. *Cereb Cortex* **13**(10): 1023-1033.

Lenochova P, Roberts SC, Havlicek J (2009). Methods of human body odor sampling: the effect of freezing. *Chem Senses* **34**(2): 127-138.

Lundqvist D, Flykt A, Ohman A (1988). *Karolinska Directed Emotional Faces* Nature Publishing Group.

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**(4): 1277-1286.

Mujica-Parodi LR, Strey HH, Frederick B, Savoy R, Cox D, Botanov Y*, et al* (2009). Chemosensory cues to conspecific emotional stress activate amygdala in humans. *PloS One* **4**(7): e6415.

Paloyelis Y, Doyle OM, Zelaya FO, Maltezos S, Williams SC, Fotopoulou A*, et al* (2016). A Spatiotemporal Profile of In Vivo Cerebral Blood Flow Changes Following Intranasal Oxytocin in Humans. *Biol Psychiatry* **79**(8): 693-705.

Prehn-Kristensen A, Wiesner C, Bergmann TO, Wolff S, Jansen O, Mehdorn HM*, et al* (2009). Induction of empathy by the smell of anxiety. *PloS One* **4**(6): e5987.

Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E*, et al* (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* **59 Suppl 20**: 22-33;quiz 34-57.

Soudry Y, Lemogne C, Malinvaud D, Consoli SM, Bonfils P (2011). Olfactory system and emotion: common substrates. *Eur Ann Otorhinolaryngol Head Neck Dis* **128**(1): 18-23.

Spengler FB, Schultz J, Scheele D, Essel M, Maier W, Heinrichs M*, et al* (2017). Kinetics and Dose Dependency of Intranasal Oxytocin Effects on Amygdala Reactivity. *Biol Psychiatry*; **82**(15): 885-894.

Spielberger CD, Gorsuch L, Lushene RE (1970). *Manual for the State-Trait Anxiety Inventory* Consulting Psychologists Press: Palo Alto, CA.

von Dawans B, Kirschbaum C, Heinrichs M (2011). The Trier Social Stress Test for Groups (TSST-G): A new research tool for controlled simultaneous social stress exposure in a group format. *Psychoneuroendocrinology* **36**(4): 514-522.

Vuilleumier P, Pourtois G (2007). Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. *Neuropsychologia* **45**(1): 174-194.

Watson D, Clark LA, Tellegen A (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol* **54**(6): 1063-1070.

Wudarczyk OA, Kohn N, Bergs R, Goerlich KS, Gur RE, Turetsky B*, et al* (2016). Chemosensory anxiety cues enhance the perception of fearful faces - An fMRI study. *Neuroimage* **143**: 214-222.

Xia M, Wang J, He Y (2013). BrainNet Viewer: a network visualization tool for human brain connectomics. *PloS one* **8**(7): e68910.

Zhou W, Chen D (2009). Fear-related chemosignals modulate recognition of fear in ambiguous facial expressions. *Psychol Sci* **20**(2): 177-183.