

Oxytocin and vasopressin effects on the neural response to social cooperation are modulated by sex in humans

Chunliang Feng · Patrick D. Hackett · Ashley C. DeMarco ·
Xu Chen · Sabrina Stair · Ebrahim Haroon · Beate Ditzen ·
Giuseppe Pagnoni · James K. Rilling

Published online: 22 November 2014
© Springer Science+Business Media New York 2014

Abstract Recent research has examined the effects of oxytocin (OT) and vasopressin (AVP) on human social behavior and brain function. However, most participants have been male, while previous research in our lab demonstrated sexually differentiated effects of OT and AVP on the neural response to reciprocated cooperation. Here we extend our previous work by significantly increasing the number of participants to enable the use of more stringent statistical thresholds that permit more precise localization of OT and AVP effects in the brain. In a double-blind, placebo-controlled study, 153 men and 151 women were randomized to receive 24 IU intranasal OT, 20 IU intranasal AVP or placebo. Afterwards, they were imaged with fMRI while playing an iterated Prisoner's Dilemma Game with same-sex partners. Sex differences were observed for effects of OT on the neural response to reciprocated cooperation, such that OT

increased the caudate/putamen response among males, whereas it decreased this response among females. Thus, 24 IU OT may increase the reward or salience of positive social interactions among men, while decreasing their reward or salience among women. Similar sex differences were also observed for AVP effects within bilateral insula and right supramarginal gyrus when a more liberal statistical threshold was employed. While our findings support previous suggestions that exogenous nonapeptides may be effective treatments for disorders such as depression and autism spectrum disorder, they caution against uniformly extending such treatments to men and women alike.

Keywords Oxytocin · Vasopressin · fMRI · Cooperation · Sex differences

Electronic supplementary material The online version of this article (doi:10.1007/s11682-014-9333-9) contains supplementary material, which is available to authorized users.

C. Feng · P. D. Hackett · X. Chen · J. K. Rilling
Department of Anthropology, Emory University, Atlanta, GA, USA

X. Chen · S. Stair · E. Haroon · B. Ditzen · J. K. Rilling
Department of Psychiatry and Behavioral Sciences, Emory University, Atlanta, GA, USA

B. Ditzen · J. K. Rilling
Center for Behavioral Neuroscience, Emory University, Atlanta, GA, USA

J. K. Rilling
Yerkes National Primate Research Center, Emory University, Atlanta, GA, USA

J. K. Rilling
Center for Translational Social Neuroscience, Emory University, Atlanta, GA, USA

A. C. DeMarco
Department of Psychology, University of Kansas, Lawrence, KS, USA

B. Ditzen
Department of Psychology, University of Zurich, Zurich, Switzerland

G. Pagnoni
Department of Neural, Biomedical, and Metabolic Sciences, University of Modena and Reggio Emilia, Modena, Italy

J. K. Rilling (✉)
Department of Anthropology, Department of Psychiatry and Behavioral Sciences, Emory University, 1557 Dickey Drive, Atlanta, GA 30322, USA
e-mail: jrilling@emory.edu

Background

Oxytocin (OT) and vasopressin (AVP) are known to modulate social behavior across a wide range of vertebrate species (Goodson and Thompson 2010). In humans, OT has been shown to support prosocial behavior in certain contexts (Heinrichs et al. 2009). For instance, intranasal administration of oxytocin increases recognition of others' facial expressions (Shahrestani et al. 2013), promotes positive communication and reduces stress hormone responses during interpersonal conflict (Ditzen et al. 2009), and enhances trusting behavior (Kosfeld et al. 2005). Fewer studies have been conducted with intranasal AVP, however recent studies have shown that intranasal AVP enhances recognition of social information and memory for social stimuli (Guastella et al. 2010b, 2011). Additionally, AVP alters friendliness judgments of same-sex, unfamiliar faces in a sexually differentiated manner (Thompson et al. 2006). Furthermore, variation in the V1a vasopressin receptor gene (AVPR1a) has been linked with variation in human social behavior (Ebstein et al. 2010; Avinun et al. 2011; Knafo et al. 2008).

Given the central role these peptides play in social behavior and cognition, there is growing interest in whether social behavioral disorders involve dysfunction in these systems, and in whether treatment with exogenous neuropeptides might improve social functioning (Meyer-Lindenberg et al. 2011; Striepens et al. 2011). Treatment with intranasal OT, which is believed to cross the blood–brain barrier and achieve access to the central nervous system (Striepens et al. 2013; Neumann et al. 2013), has been investigated in a variety of psychiatric disorders, including autism spectrum disorder (ASD), depression, schizophrenia, and social anxiety disorder (Striepens et al. 2011; Macdonald and Feifel 2013). In individuals with ASD, intranasal OT increases gaze to the eye region of the face, improves accuracy of mental state attribution from the eye region of the face, and biases social interactions towards more cooperative social partners (Guastella et al. 2010a; Andari et al. 2010). A recent fMRI study in children with high-functioning ASD provided a potential mechanism for these effects, showing that OT increased activity during social judgments and decreased activity during nonsocial judgments in the striatum and nucleus accumbens, regions involved in salience and reward processing (Gordon et al. 2013). OT has also been shown to normalize the amygdala hyperactivity characteristic of social anxiety disorder (Labuschagne et al. 2010; Dohia et al. 2014). Against the back-drop of these promising clinical studies is accumulating evidence for sex differences in the effects of OT and AVP in both non-human animals and humans (Albers 2012; Veenema et al. 2013; Dumais et al. 2013; Hoge et al. 2014; Thompson et al. 2006; Lynn et al. 2014; Campbell et al. 2014). Some of these studies actually report opposite effects of OT and AVP on the social behavior of males and females, raising the possibility that

treatment with neuropeptides may be appropriate for patients of one sex but not the other.

Previously, we reported pronounced sex differences in the effects of OT and AVP on the neural response to positive social interactions (cooperative exchanges in the iterated Prisoner's Dilemma (PD) game) in a sample of 91 males and 87 females (Rilling et al. 2014). Specifically, we showed that the contrast between males and females for OT effects yielded a single large cluster of activation (1,733 voxels) that spanned multiple brain regions including the amygdala, caudate/putamen, insula, basal forebrain, and hippocampus, among others. For AVP, the same contrast yielded just two large clusters: one of them (3,333 voxels) spanned the amygdala, striatum, insula, thalamus, and cingulate cortex, among others. The other cluster (615) included the postcentral gyrus, precentral gyrus, supramarginal gyrus and transverse temporal gyrus. However, as discussed in a recent methodological paper (Woo et al. 2014), the cluster-extent based thresholding we employed (voxel-wise $p < 0.05$ in conjunction with cluster-wise $p < 0.05$ family-wise error [FWE] corrected), although an acceptable means of correcting for multiple comparisons, does not permit identification of specific brain structures within those large clusters that constitute true positive activations. The current study builds on our prior report by expanding the sample to include an additional 62 males and 64 females and applies the more stringent statistical threshold recommended by (Woo et al. 2014) (voxel-wise $p < 0.001$ in conjunction with cluster-wise $p < 0.05$ FWE corrected) to better isolate true positive activations. With these methods, we find that sex differences in OT effects could be specifically localized to the bilateral caudate/putamen, superior medial frontal cortex and frontal pole, where OT augmented men's neural response to cooperation, rendering men similar to women in the placebo condition and OT attenuated women's neural response to cooperation, rendering women more similar to men in the placebo group. With a slightly less stringent threshold (voxel-wise $p < 0.005$ in conjunction with cluster-wise $p < 0.05$ FWE corrected) we also show similar sex differences in AVP effects on neural response to reciprocated cooperation within bilateral insula and left supramarginal gyrus.

Methods

Subjects

153 men and 151 women from the Emory University community between the ages of 18 and 22 (mean age for men = 20.7 years; mean age for women = 20.5 years) were randomized to receive intranasal OT (50 men and 50 women), intranasal AVP (49 men and 51 women), or intranasal placebo (54 men and 50 women). This sample includes and expands upon a sample on which we have previously published (Rilling et al.

2012, 2014). The earlier sample consisted of 91 males (OT $n=27$, AVP $n=27$, PBO $n=37$) and 87 females (OT $n=29$, AVP $n=29$, PBO $n=29$). Here, we add an additional 62 males and 64 females. All potential subjects completed a full medical history questionnaire. Exclusion criteria are provided in the Online Resource 1. All subjects gave written informed consent, and the study was approved by the Emory University Institutional Review Board and the U.S. Food and Drug Administration. Eight men (OT $n=3$, AVP $n=2$, and placebo $n=3$) and two women (OT $n=2$) were excluded from the neuroimaging analysis due to excessive motion (>1.5 mm) ($n=7$), missing data ($n=2$) or to abnormal brain anatomy ($n=1$). Also, four men (OT $n=2$, AVP $n=1$, and placebo $n=1$) and eleven women (OT $n=4$, AVP $n=4$, and placebo $n=3$) had only partial data due to excessive motion during one portion of the scan.

Behavioral procedures

Administration of OT, AVP or placebo

Both experimenters and subjects were blind to the treatment subjects received. All solutions were administered intranasally (preparation of OT, AVP and placebo is described in Online Resource 1). Following previous studies that have reported social cognitive effects of intranasal OT and AVP administration (Macdonald and Feifel 2013; Thompson et al. 2006), the OT group self-administered 24 IU oxytocin (Syntocinon-Spray, Novartis), and the AVP group self-administered 20 IU of AVP (American Reagent Laboratories, Shirley, NY, USA). In each case, this required 10 nasal puffs to administer 1 ml of solution. The placebo group self-administered 10 nasal puffs of either OT placebo or AVP placebo (both including all ingredients, i.e., preservatives, without the active pharmacological substance). Half of the placebo subjects received OT placebo and half received AVP placebo. Subjects were instructed to place the nasal applicator in one nostril and depress the lever until they felt a mist of spray in the nostril, to then breathe in deeply through the nose, and afterwards to place the applicator in the other nostril and repeat the process.

Prisoner's Dilemma task

The iterated Prisoner's Dilemma (PD) game is a model for relationships based on reciprocal altruism (Fig. 1). In the game, two players choose to either cooperate or defect and receive a payoff that depends upon the interaction of their respective choices. The game version used in the current study is a sequential-choice PD game, in which player 1 chooses and player 2 is then able to view player 1's choice before making his/her own choice. Each of the four outcomes is associated with a different payoff. Player cooperation followed by partner cooperation (CC) pays \$2 to both player and partner,

player cooperation followed by partner defection (CD) pays \$0 to the player and \$3 to the partner, player defection followed by partner defection (DD) pays \$1 to both player and partner, and player defection followed by partner cooperation (DC) pays \$3 to the player and \$0 to the partner. While being imaged with fMRI, subjects played 30 rounds of a sequential-choice, iterated Prisoner's Dilemma game in each of four sessions. For two sessions, subjects were told they were playing with the human partners they were introduced to (confederate introductions are detailed in Online Resource 1). For the other two sessions, subjects were told that they were playing with a computer partner. In reality, subjects were always playing with a pre-programmed computer algorithm (see Online Resource 1 for details) in all four sessions. For both human and computer partners, in one of the two sessions, subjects played in the role of first mover (player 1) and their partner played in the role of second mover (player 2). In the second session, roles were reversed (the counterbalancing of human and computer sessions is described in the Online Resource 1). Subjects were compensated with a total of approximately \$120; the exact amount was obtained by multiplying the total earnings across all four runs of the PD Game by 2/3.

Behavioral analysis

In this manuscript, statistical analyses are limited to player 1 data. Outcome frequencies and transition probabilities were compared across partner types (human, computer) using paired t -tests and across treatment groups (OT, AVP, placebo) using two sample t -tests. We use t -tests rather than ANOVA to evaluate treatment group effects because we are specifically interested in whether each drug group (OT, AVP) differs from placebo rather than whether there is a difference among all three groups. Furthermore, OT, AVP, and placebo cannot be considered levels of a single factor in the way that, for example, low, medium, and high body weight is. Outcomes (Supplementary Table 1 in the Online Resource 2) included CC (player cooperate followed by partner cooperate), CD (player cooperate followed by partner defect), DC (player defect followed by partner cooperate) and DD (player defect followed by partner defect), while transition probabilities were the probability of cooperating conditional on the outcome of the previous round. Therefore, there were four different transition probabilities: the probability of cooperating after a CC outcome (pC/CC), after a CD outcome (pC/CD), after a DC outcome (pC/DC) and after a DD outcome (pC/DD). In the current study, we focused specifically on the behavioral and neural response to CC outcomes for two reasons: First, the current study aimed to extend and refine our previous study in which CC outcomes were the focus (Rilling et al. 2014). Second, OT has been shown to promote prosocial behavior in both humans and non-human animals. OT facilitates

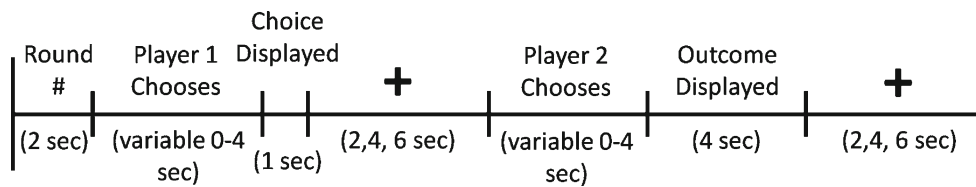


Fig. 1 Timeline of Prisoner's Dilemma task. E-prime software (Psychology Software Tools, Pittsburgh, PA, USA) was used for stimulus presentation. Stimuli were projected onto a screen that subjects could view through a mirror mounted on the head coil. Subjects' responses were recorded using a response box. At the beginning of each round, the round number and partner's photo were displayed for 2 s. Player 1 then had 4 s to choose to cooperate or defect. Players were informed that if they did not decide within this 4 s interval, their response would default to defection. Player 1's choice was immediately revealed to player 2 and

displayed for 1 s. A variable length fixation epoch of either 2, 4, or 6 s followed. Afterwards, player 2 had 4 s to cooperate or defect. Once player 2 decided, the outcome of the round was displayed for 4 s. Finally, the trial concluded with another variable length fixation epoch of either 2, 4 or 6 s. Trials were approximately 20 s long. Five null trials were interspersed among 30 PD trials in each session. Null trials consisted of 14 s of fixation. One session lasted approximately 12 min. All four sessions lasted about 48 min

maternal caregiving and female partner-preference among rodents (Liu and Wang 2003; Ross and Young 2009). In humans, intranasal OT enhances trust behavior (Kosfeld et al. 2005) and promotes positive communication and reduces stress hormone responses during interpersonal conflict (Ditzen et al. 2009). These findings suggest that OT may enhance sensitivity to reward-relevant features of human social interactions (Marsh et al. 2010; Rilling et al. 2012). Similarly, many studies have demonstrated that AVP can facilitate social affiliation in some non-human animals in some contexts (Goodson and Thompson 2010). Therefore, we only report results on the frequency of CC outcomes and pC/CC here. Sex differences in drug effects were evaluated using two-way ANOVA.

Neuroimaging procedures

Anatomical image acquisition

Subjects were positioned head first in the supine position inside the scanner (Siemens Trio 3 T), with padded head restraint to minimize head motion during scanning. Each scanning session began with a 15 s scout, followed by a 5 min T1-weighted MPRAGE 3d scan that was acquired in the sagittal plane and accelerated by generalized auto-calibrating partially parallel acquisitions (GRAPPA) with a factor of 2 (TR=2,600 ms, TE=3.02 ms, matrix=256×256×176, FOV=256 mm×256 mm×176 mm, slice thickness=1.00 mm, gap=0 mm).

fMRI image acquisition

Subjects were imaged while playing the PD game. Functional scans used an EPI sequence with the following parameters: TR=2,000 ms, TE=28 ms, matrix=64×64, FOV=224 mm, slice thickness=2.5 mm, 34 axial slices with a slice gap of 1.05 mm. TE was minimally decreased from the typical value (32 ms) in order to reduce

magnetic susceptibility artifact in the orbitofrontal region. The duration of each EPI scan was about 12 min (30 PD round×~20 s per round, plus five null trials×14 s per trial). After each of the four sessions, while still in the scanner, subjects rated their emotional reaction to the four PD game outcomes (CC, CD, DC, and DD). Seven-point Likert scales were used to rate the following emotions or feelings: afraid, angry, happy, guilty, disappointed, and relieved.

MRI image analysis

Image processing was conducted with FEAT (fMRI Expert Analysis Tool) version 6.00, part of FSL (FMRIB's Software Library, <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>). Preprocessing involved motion correction using MCFLIRT (Jenkinson et al. 2002), slice timing correction using Fourier-space time-series phase-shifting, non-brain removal using BET (Smith 2002), spatial smoothing using a Gaussian kernel of FWHM 5 mm, grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor, and highpass temporal filtering (Gaussian-weighted least-squares straight line fitting, with cut-off =100.0 s). Registration to MNI space via corresponding extracted T1 brain was carried out using Boundary-Based-Registration (Greve and Fischl 2009). Time-series statistical analysis was carried out using FILM with local autocorrelation correction (Woolrich et al. 2001).

For the player 1 runs, a separate general linear model (GLM) was defined for each subject that examined the neural response to both the epoch in which the choice to cooperate or defect was made, as well as to the epoch in which the trial outcome was revealed. More specifically, the following regressors were defined for each subject in the role of player 1: (1) the beginning epoch when round number and the partner's face or a picture of computer was displayed, (2) the choice epoch when the subject chose to

cooperate (choice C), (3) the choice epoch when the subject chose to defect (Choice D), (4) CC outcomes, (5) CD outcomes, (6) DC outcomes, and (7) DD outcomes. These regressors were specified separately for runs with human and computer partners, resulting in a total of 14 distinct regressors per subject. Parameter estimates for CC outcomes with human and computer partners were computed at every voxel within the brain. Higher level (group level) analyses were carried out using an Ordinary Least Square (OLS) model in FEAT. At each voxel, the interaction between sex and drug treatment was calculated by specifying the contrast 1–1 –1 1 (male with peptide treatment, male with placebo treatment, female with peptide treatment, female with placebo treatment) separately for OT and AVP. *T*-statistics were converted to *Z*-statistics by equating probabilities under the tails of the *T* and *Z* distributions. The *Z*-statistic (“Gaussianised *T*”) images were thresholded using clusters determined by $Z > 3.3$ (voxel-wise 2-tailed $p < 0.001$), and a family-wise error (FWE)-corrected cluster significance threshold of $p < 0.05$ was applied to the supra-threshold clusters. This is the multiple comparison correction threshold recommended by Woo et al. (2014) to ensure better localization of the significant activations. To further examine the gender by drug treatment interaction effect, ROI analyses were conducted to reveal the simple effects of OT or AVP in males and females within areas showing a significant interaction effect, which were Bonferroni corrected for the total number of tests (Harlé et al. 2012; Güroğlu et al. 2011).

Results

Behavioral results

Each subject played 30 rounds of the PD game with both an assumed human and a computer partner in the role of first mover (player 1). Collapsing across all drug treatment conditions, the average number of experienced CC outcomes was 11.21 ± 0.47 in males and 11.36 ± 0.45 in females for human partners; 11.31 ± 0.47 in males and 11.91 ± 0.44 in females for computer partners. Two-way ANOVA did not reveal any effects of drug treatment or interaction with sex on either CC outcome frequency or the probability of cooperating after a CC outcome in the previous round (pC/CC), for either male or female subjects interacting with either presumed human or computer partners (Fig. 2). However, within the AVP group, pC/CC with human partners was higher for males than females ($n=99$, two sample *t*-test, $t_{97}=2.00$, $p < 0.05$, uncorrected).

Neuroimaging results

Sex differences in OT effects on the BOLD response to reciprocated cooperation

In our previous report, which employed a voxel-wise threshold of $p < 0.05$ along with a cluster threshold $p < 0.05$ FWE corrected (Rilling et al. 2014), the interaction between sex and OT treatment ($[(\text{male OT group} - \text{male placebo group}) - (\text{female OT group} - \text{female placebo group})]$) yielded a single large cluster of activation (1,733 voxels) that spanned multiple brain regions, including the amygdala, caudate/putamen, insula, basal forebrain, and hippocampus, among others. In the current study, the interaction between sex and drug treatment ($[(\text{male OT group} - \text{male placebo group}) - (\text{female OT group} - \text{female placebo group})]$) yielded more focal activation in four distinct regions: the left ($-18/18/6$ mm, cluster size=255, $Z=4.21$) and right ($22/20/-4$ mm, cluster size=251, $Z=4.35$) caudate/putamen, the right frontal pole ($18/60/20$ mm, cluster size=169, $Z=4.20$) and the left medial part of superior medial frontal cortex ($-4/12/56$ mm, cluster size=121, $Z=4.20$) (Table 1 & Fig. 3a). Within these regions, OT increased activation relative to placebo in males ($p < 0.05$, Bonferroni corrected), while decreasing activation relative to placebo in females ($p < 0.05$, Bonferroni corrected) (Fig. 3b & Supplementary Fig. 1 in Online Resource 2). In contrast to these findings with human partners, a test for interaction of gender and drug treatment did not reveal any significant activations in response to CC outcomes with computer partners, even with a more liberal voxel-wise threshold of $p < 0.005$.

Sex differences in AVP effects on the BOLD response to reciprocated cooperation

In our previous report with voxel-wise threshold of $p < 0.05$ along with the cluster threshold $p < 0.05$ FWE corrected (Rilling et al. 2014), the interaction between sex and AVP treatment ($[(\text{male AVP group} - \text{male placebo group}) - (\text{female AVP group} - \text{female placebo group})]$) yielded activation in two large clusters: one of them (3,333 voxels) spanned multiple brain regions including the amygdala, striatum, insula, thalamus, and cingulate cortex, among others. The other cluster (615) included the postcentral gyrus, precentral gyrus, supramarginal gyrus and transverse temporal gyrus. In the current study, a whole-brain analyses did not reveal sex differences in AVP effects on neural response to CC outcomes with human partners, using the recommended voxel-wise threshold of $p < 0.001$ and cluster extent threshold of $p < 0.05$ FWE corrected. However, with a more liberal voxel-wise threshold of $p < 0.005$ along with the cluster threshold $p < 0.05$ FWE corrected, the interaction of gender and drug treatment ($[(\text{male AVP group} - \text{male placebo group}) - (\text{female AVP group} - \text{female placebo group})]$) yielded

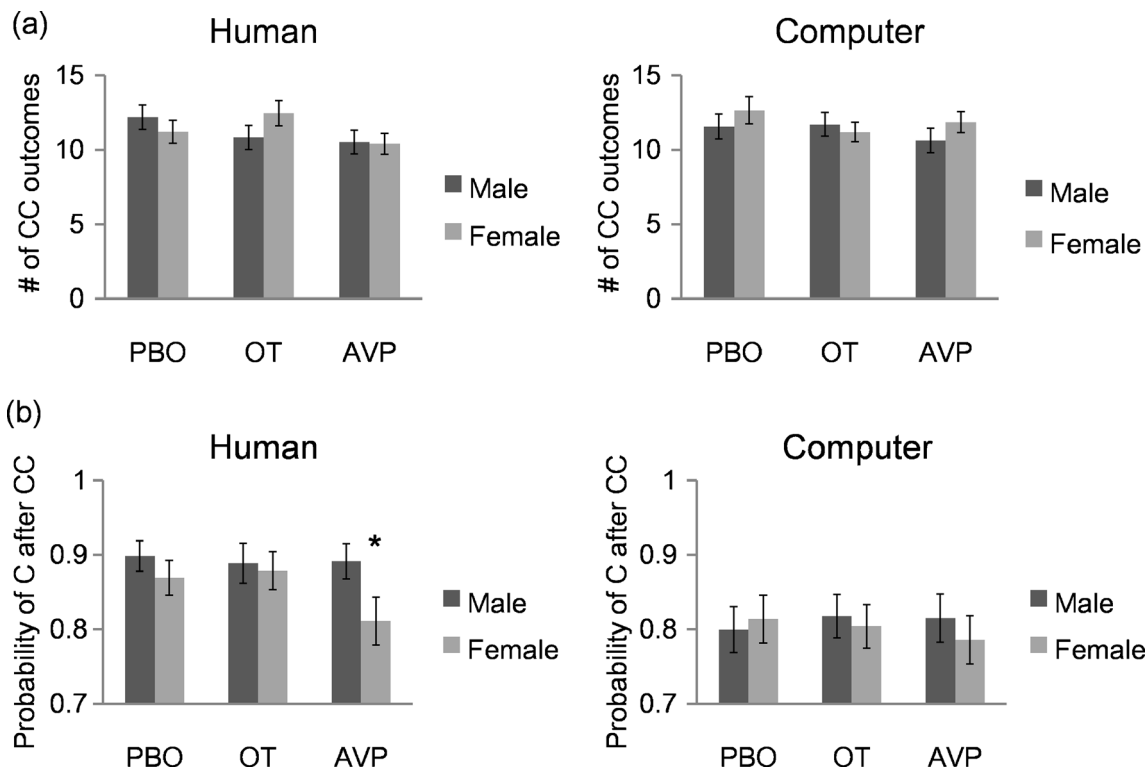


Fig. 2 Behavioral results. **a** CC outcome frequency with human (*left*) or computer (*right*) partners as player 1, as a function of sex and drug treatment groups. No effects of sex or drug were revealed. **b** Probability of cooperating after a CC outcomes (pC/CC) with human (*left*) or

computer (*right*) partners as player 1, as a function of sex and drug treatment groups. Within the AVP group, pC/CC was higher for males than females. Error bars represent one standard error. * $p < 0.05$ (uncorrected). PBO placebo, OT oxytocin, AVP vasopressin

activation in the left insula (−30/12/10 mm, cluster size=654, $Z=4.57$), right insula (42/−6/−4 mm, cluster size=260, $Z=4.04$) and left supramarginal gyrus (−48/−40/42 mm, cluster size=352, $Z=3.88$) (Supplementary Table 2 & Supplementary Fig. 2a in Online Resource 2). Within each of these regions, AVP augmented activation relative to placebo in males

($p < 0.05$, Bonferroni corrected), but attenuated activation relative to placebo in females ($p < 0.05$, Bonferroni corrected) (Supplementary Fig. 2b in Online Resource 2). Even with the liberal voxel-wise threshold of $p < 0.005$, the test for interaction did not reveal any significant activations in response to CC outcomes with computer partners.

Table 1 Brain regions where the BOLD response to reciprocated cooperation (CCh) with human partners was modulated by the interaction between OT treatment and sex (*male: OT-placebo >female:*

OT-placebo). Whole brain random effect analysis; voxel-wise $p < 0.001$ in conjunction with cluster-wise $p < 0.05$ FWE-corrected

Brain Regions	MNI Coordination of Local Maxima (mm)			Local Maxima Z	Cluster Size (voxel)
	x	y	z		
Left Caudatae/Putamen/Planum Polare	−18	18	6	4.21	255
Left Caudate	−18	18	6	4.21	
Left Putamen	−28	6	8	4.09	
Left Planum Polare	−26	4	−12	3.368	
Right Putamen/Caudate	22	20	−4	4.35	251
Right Putamen	22	20	−4	4.35	
Right Caudate	18	20	12	4.25	
Right Frontal Pole	18	60	20	4.20	169
Right Frontal Pole	18	60	20	4.20	
Superior Medial Frontal Gyrus	−4	12	56	4.20	121
Left Superior Medial Frontal Gyrus	−4	12	56	4.20	
Right Superior Medial Frontal Gyrus	2	18	66	3.68	

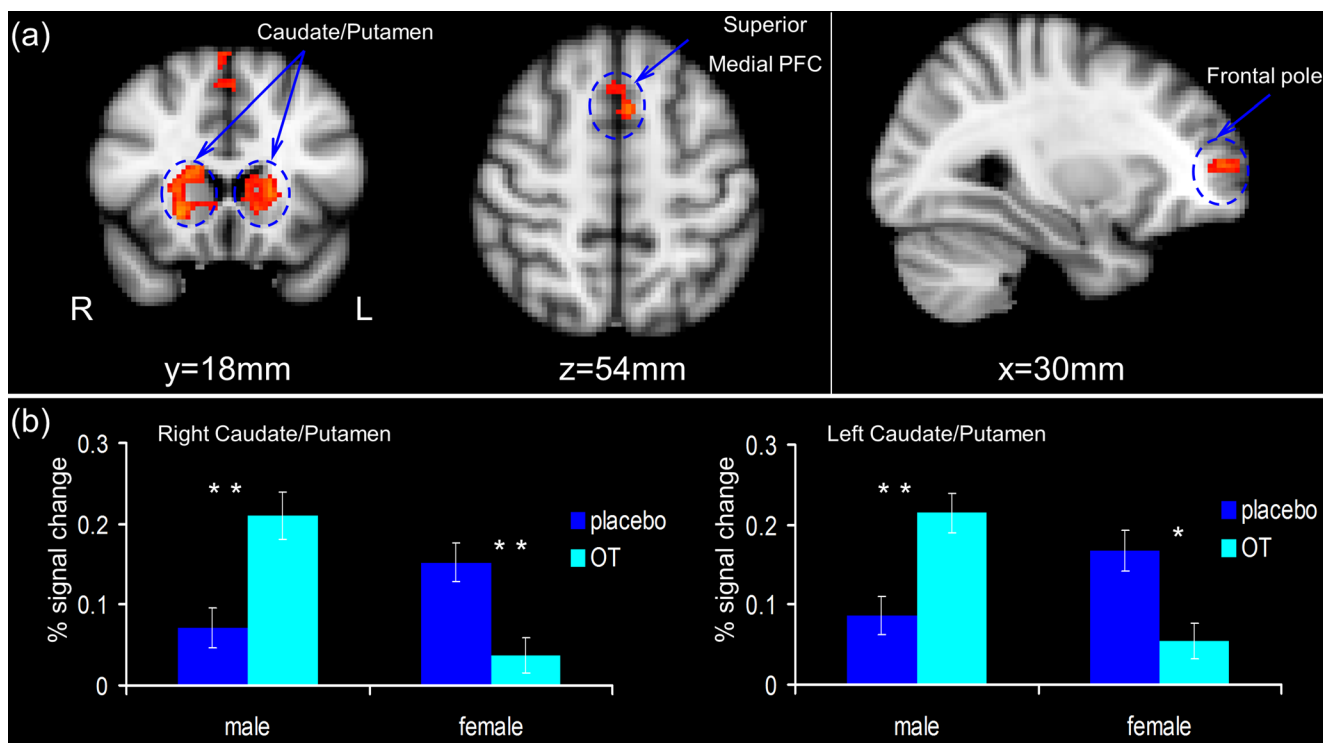


Fig. 3 Sex differences in OT effects on brain activations in response to CC outcomes with human partners. **a** Activations revealed by whole brain analysis, voxel-wise threshold of $p < 0.001$ in conjunction with cluster-wise threshold of $p < 0.05$ FWE-corrected; **b** ROI plot from **a**) illustrating the caudate/putamen response as a function of sex and drug treatment. OT

augmented neural response to CC outcomes from human partners among males, whereas the opposite was found for females. Error bars represent one standard error. * $p < 0.05$; ** $p < 0.005$ (Bonferroni corrected). R right, L left

Discussion

The goal of the current study was to extend and refine our previous findings by expanding our sample size and by applying a more stringent statistical threshold for cluster-extent based multiple comparison correction with the aim of achieving a more precise understanding of how OT and AVP differentially modulate the neural response to social cooperation in men and women. Sex differences in OT effects were localized to caudate/putamen. Although these brain regions were also identified in our previous report, they were part of a large cluster of activation that spanned multiple brain regions. With the statistical thresholds we employed previously, it was not possible to discern the precise anatomical focus of true positive activation within that large cluster (Woo et al. 2014). Here, with more stringent statistical methods, we localize that effect to the caudate/putamen. This implies that the activations previously observed in amygdala, insula, basal forebrain, and hippocampus, among other regions, are unlikely to be true positive activations, at least with the level of sensitivity afforded by our current sample size. The current study also revealed two activations that were not found in our earlier analysis: the frontal pole and superior medial prefrontal gyrus. In caudate/putamen, frontal pole and medial prefrontal gyrus identified in the current study, OT increased the BOLD fMRI

response to reciprocated cooperation from human partners in males, whereas OT decreased the BOLD fMRI response to reciprocated cooperation in females. Given the role of caudate nucleus in registering social prediction errors that guide decisions about reciprocity in Trust or PD games (Rilling et al. 2002; King-Casas et al. 2005), the current findings may indicate that OT enhances the reward from reciprocated cooperation and/or facilitates learning that another person can be trusted to reciprocate cooperation among males, whereas OT may have the opposite effects among females. Research with non-human animals has established that OT and dopamine interact in the ventral striatum to facilitate social bonding and attachment (Skuse and Gallagher 2009; Young et al. 2005). Our finding that OT augments the striatal response to reciprocated cooperation among men suggests that similar mechanisms may be operational in humans. The striatum is also involved in salience detection (Zink et al. 2004), so our results might also imply that OT enhances the salience of positive social interactions among men, while decreasing their salience among women.

Sex differences in OT effects were also observed in medial prefrontal cortex and frontal pole. Medial prefrontal cortex is associated with overtly thinking about the internal mental states of others (i.e., mentalizing) during social interactions (Amodio and Frith 2006; Rilling et al. 2004). Therefore,

effects of OT on the medial prefrontal cortex are consistent with previous evidence that OT modulates “mind-reading” in humans (Domes et al. 2007). These OT effects are also consistent with evidence that OT mitigates sociocommunicational deficits of ASD by increasing activity in the medial prefrontal cortex (Watanabe et al. 2014). The frontal pole is hypothesized to represent the value of future or abstract outcomes in social decision making (Rilling and Sanfey 2011), such that the activation of frontal pole is observed when participants decide to trust others (Krueger et al. 2008). In this regard, effects of OT on frontal pole is consistent with previous finding that OT influences trust behaviors in humans (Kosfeld et al. 2005). Taken together, OT effects on medial prefrontal cortex and frontal pole in the current study may reflect the attribution of partner intention from reciprocated cooperation and the evaluation of possible options for choice in the next round.

Similar sex differences in effects of AVP were observed in bilateral insula and right SMG. In our previous study, these brain regions were also identified as part of large clusters spanning multiple brain regions. In addition, the activity previously observed in the amygdala, striatum, thalamus, and cingulate cortex, among other regions, was not observed in the current study with more stringent statistical methods, implying that activations in these brain regions are unlikely to be true positive activations, at least with the level of sensitivity afforded by our current sample size. In bilateral insula and right SMG, AVP enhanced activation relative to placebo among men, while decreasing activation relative to placebo among women. Similar to the striatum, the insula has also been implicated in salience detection (Menon and Uddin 2010), raising the possibility that AVP is increasing the salience of positive social interactions among men, while decreasing their salience among women. Interestingly, these sex differences in AVP effects on brain activity were paralleled by a lower probability of persisting with cooperation (pC/CC) among women compared with men in the AVP group. Thus, AVP may be lowering this probability by decreasing the salience of positive social interactions among women.

Sex differences in OT and AVP effects were in the same direction. That is, both neuropeptides enhanced the neural response to CC outcomes in males but decreased the neural response to CC outcomes in females. This suggests that OT and AVP modulate social cognition in a similar way, at least in the case of positive social interactions (Goodson and Thompson 2010). However, this statement requires qualification for two reasons. First, OT and AVP effects were found in different brain regions, implying differential neural mechanisms for these two neuropeptides. Second, some studies have shown opposing effects of OT and AVP in other contexts, such as the anxiolytic effects of OT and anxiogenic effects of AVP in response to stressful social interactions (Heinrichs et al. 2009).

Our study is quite large by fMRI standards. Increasing our sample size from 178 to 304 allowed us to more precisely identify the location of sex differences in OT and AVP effects. It also revealed additional activations that were not present in our earlier sample. However, it is difficult to conclude from this study that a particular sample size is needed to reveal OT and AVP effects on brain activation in general since the required sample size is likely to depend on the task that is used (e.g., simple perceptual task vs. complex social-interaction task) as well as the experimental design (e.g., within-subject design vs. between-subject design) employed.

Consistent with our findings here, accumulating studies have reported sexually differentiated effects of AVP and OT on social cognition and behavior in both non-human animals and humans (Albers 2012; Veenema et al. 2013; Dumais et al. 2013; Hoge et al. 2014; Lynn et al. 2014; Bales and Carter 2003; Fischer-Shofty et al. 2013). At baseline in the placebo condition, women had a stronger caudate/putamen response to reciprocated cooperation than men did (right caudate/putamen $t_{92}=2.31$, $p=0.02$; left caudate/putamen $t_{92}=2.34$, $p=0.02$). This coupled with evidence that women have higher baseline OT levels in cerebrospinal fluid (CSF) (Altemus et al. 1999), suggests the possibility of an inverted-U shaped dose response function between brain OT levels and neural activity, whereby raising brain OT levels in men would augment the caudate/putamen response, moving them closer to the maximum level of activity. On the other hand, raising OT levels in women might displace them to the right of the maximum, decreasing brain activity. This hypothesis is supported by recent studies investigating the dose–response properties of intranasal OT among human males. In particular, the 24 IU dose of intranasal OT, as compared with 48 IU dose and placebo, reduced men’s cortisol response to physical stress, enhanced the number of specific autobiographical memories retrieved, and promoted the retrieval of social affiliation memories associated with more positive feelings, whereas no difference was observed between 48 IU dose and placebo on these measures (Cardoso et al. 2013, 2014). However, the inverted U-shaped interpretation of our OT findings is obviously speculative given that we did not measure baseline and post-drug levels of OT in CSF. This hypothesis should be further evaluated in future experiments. In contrast to OT, baseline CSF AVP levels are reported to be similar in men and women (Coccaro et al. 1998; Sørensen et al. 1985), implying that an inverted U shaped response curve may not be the explanation for our AVP results. Potential alternative explanations include possible sex differences in AVPR1a receptor expression, which has not been systematically investigated in humans, and interactions with steroid hormones, levels of which differ between men and women.

Although none of the brain regions where we found modulatory effects of OT or AVP are reported to have high densities of OT or AVP receptors (Loup et al. 1991; Boccia et al.

2013), OT and AVP effects in these regions could be accounted for by downstream effects of OT or AVP binding to other brain structures.

Despite these effects of intranasal OT on the BOLD fMRI response to reciprocated cooperation, intranasal OT had no effect on the behavioral response to this outcome (pC/CC). Recent studies show that the effects of intranasal OT on cooperative behavior depend on the participant's attachment style (Bartz et al. 2011). For example, intranasal OT increases cooperation among men who are high rather than low in attachment avoidance (De Dreu 2012). We did not characterize the attachment style of the participants in our study, so we are unable to evaluate the potential influence of attachment style on neural and behavioral responsiveness to OT.

Recently, there has been growing interest in the potential use of OT as a treatment for a variety of psychiatric disorders, including ASD, schizophrenia, anxiety disorders and depression (Macdonald and Feifel 2013; Striepens et al. 2011). Our findings may be particularly relevant for ASD and depression. By enhancing the reward from positive social interactions, OT could potentially increase the motivation of males with ASD to engage with others, and a recent study even showed that intranasal OT increases striatal activation during social judgments in children with high-functioning ASD (Gordon et al. 2013). Critically, we find that OT only has these properties in men, suggesting that the clinical efficacy of 24 IU intranasal OT may be limited to males and urging caution against uniformly extending these treatments to men and women alike. In future studies, it will be important to evaluate the effects of both OT and AVP across a range of doses to establish the dose–response relationship for the effects observed here. It will also be important to study men and women interacting with opposite-sex partners to determine if the effects observed here depend simply on the sex of the player or rather the interaction of player and partner sex (Preckel et al. 2014).

Acknowledgments We thank Susan Rogers, Jianguo Xu and Larry Young for assistance with various aspects of this study. This study was supported by National Institute of Mental Health [grant number R01 MH084068-01A1] and the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number UL1TR000454. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest Chunliang Feng, Patrick D. Hackett, Ashley C. DeMarco, Xu Chen, Sabrina Stair, Ebrahim Haroon, Beate Ditzen, Giuseppe Pagnoni and James K. Rilling declare that they have no conflict of interest.

Informed consent All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, and the applicable revisions at the time of the investigation. Informed consent was obtained from all subjects for being included in the study.

References

- Albers, H. E. (2012). The regulation of social recognition, social communication and aggression: vasopressin in the social behavior neural network. *Hormones and Behavior*, *61*(3), 283–292.
- Altemus, M., Jacobson, K. R., Debellis, M., Kling, M., Pigott, T., Murphy, D. L., et al. (1999). Normal CSF oxytocin and NPY levels in OCD. *Biological Psychiatry*, *45*(7), 931–933.
- Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*(4), 268–277.
- Andari, E., Duhamel, J.-R., Zalla, T., Herbrecht, E., Leboyer, M., & Sirigu, A. (2010). Promoting social behavior with oxytocin in high-functioning autism spectrum disorders. *Proceedings of the National Academy of Sciences*, *107*(9), 4389–4394.
- Avinun, R., Israel, S., Shalev, I., Gritsenko, I., Bornstein, G., Ebstein, R. P., et al. (2011). AVPR1A variant associated with preschoolers' lower altruistic behavior. *PloS One*, *6*(9), e25274.
- Bales, K. L., & Carter, C. S. (2003). Sex differences and developmental effects of oxytocin on aggression and social behavior in prairie voles (*Microtus ochrogaster*). *Hormones and Behavior*, *44*(3), 178–184.
- Bartz, J. A., Zaki, J., Bolger, N., & Ochsner, K. N. (2011). Social effects of oxytocin in humans: context and person matter. *Trends in Cognitive Sciences*, *15*(7), 301–309.
- Boccia, M., Petrusz, P., Suzuki, K., Marson, L., & Pedersen, C. (2013). Immunohistochemical localization of oxytocin receptors in human brain. *Neuroscience*, *253*, 155–164.
- Campbell, A., Ruffman, T., Murray, J., & Glue, P. (2014). Oxytocin improves emotion recognition for older males. *Neurobiology of Aging* (ahead-of-print), 1–13.
- Cardoso, C., Ellenbogen, M. A., Orlando, M. A., Bacon, S. L., & Joobar, R. (2013). Intranasal oxytocin attenuates the cortisol response to physical stress: a dose–response study. *Psychoneuroendocrinology*, *38*(3), 399–407.
- Cardoso, C., Orlando, M. A., Brown, C. A., & Ellenbogen, M. A. (2014). Oxytocin and enhancement of the positive valence of social affiliation memories: An autobiographical memory study. *Social neuroscience* (ahead-of-print), 1–10.
- Coccaro, E. F., Kavoussi, R. J., Hauger, R. L., Cooper, T. B., & Ferris, C. F. (1998). Cerebrospinal fluid vasopressin levels: correlates with aggression and serotonin function in personality-disordered subjects. *Archives of General Psychiatry*, *55*(8), 708–714.
- De Dreu, C. K. (2012). Oxytocin modulates the link between adult attachment and cooperation through reduced betrayal aversion. *Psychoneuroendocrinology*, *37*(7), 871–880.
- Ditzen, B., Schaer, M., Gabriel, B., Bodenmann, G., Ehlert, U., & Heinrichs, M. (2009). Intranasal oxytocin increases positive communication and reduces cortisol levels during couple conflict. *Biological Psychiatry*, *65*(9), 728–731.
- Dodhia, S., Hosanagar, A., Fitzgerald, D. A., Labuschagne, I., Wood, A. G., Nathan, P. J., et al. (2014). Modulation of Resting-State Amygdala-Frontal Functional Connectivity by Oxytocin in Generalized Social Anxiety Disorder. *Neuropsychopharmacology* (ahead of print).
- Domes, G., Heinrichs, M., Michel, A., Berger, C., & Herpertz, S. C. (2007). Oxytocin improves “mind-reading” in humans. *Biological Psychiatry*, *61*(6), 731–733.
- Dumais, K. M., Bredewold, R., Mayer, T. E., & Veenema, A. H. (2013). Sex differences in oxytocin receptor binding in forebrain regions: correlations with social interest in brain region-and sex-specific ways. *Hormones and Behavior*, *64*(4), 693–701.
- Ebstein, R. P., Israel, S., Chew, S. H., Zhong, S., & Knafo, A. (2010). Genetics of human social behavior. *Neuron*, *65*(6), 831–844.
- Fischer-Shofty, M., Levkovitz, Y., & Shamay-Tsoory, S. G. (2013). Oxytocin facilitates accurate perception of competition in men and

- kinship in women. *Social Cognitive and Affective Neuroscience*, 8(3), 313–317.
- Goodson, J. L., & Thompson, R. R. (2010). Nonapeptide mechanisms of social cognition, behavior and species-specific social systems. *Current Opinion in Neurobiology*, 20(6), 784–794.
- Gordon, I., Vander Wyk, B. C., Bennett, R. H., Cordeaux, C., Lucas, M. V., Eilbott, J. A., et al. (2013). Oxytocin enhances brain function in children with autism. *Proceedings of the National Academy of Sciences*, 110(52), 20953–20958.
- Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, 48(1), 63–72.
- Guastella, A. J., Einfeld, S. L., Gray, K. M., Rinehart, N. J., Tonge, B. J., Lambert, T. J., et al. (2010a). Intranasal oxytocin improves emotion recognition for youth with autism spectrum disorders. *Biological Psychiatry*, 67(7), 692–694.
- Guastella, A. J., Kenyon, A. R., Alvares, G. A., Carson, D. S., & Hickie, I. B. (2010b). Intranasal arginine vasopressin enhances the encoding of happy and angry faces in humans. *Biological Psychiatry*, 67(12), 1220–1222.
- Guastella, A. J., Kenyon, A. R., Unkelbach, C., Alvares, G. A., & Hickie, I. B. (2011). Arginine Vasopressin selectively enhances recognition of sexual cues in male humans. *Psychoneuroendocrinology*, 36(2), 294–297.
- Güroğlu, B., van den Bos, W., van Dijk, E., Rombouts, S. A., & Crone, E. A. (2011). Dissociable brain networks involved in development of fairness considerations: understanding intentionality behind unfairness. *NeuroImage*, 57(2), 634–641.
- Harlé, K. M., Chang, L. J., van't Wout, M., & Sanfey, A. G. (2012). The neural mechanisms of affect infusion in social economic decision-making: a mediating role of the anterior insula. *NeuroImage*, 61(1), 32–40.
- Heinrichs, M., von Dawans, B., & Domes, G. (2009). Oxytocin, vasopressin, and human social behavior. *Frontiers in Neuroendocrinology*, 30(4), 548–557.
- Hoge, E. A., Anderson, E., Lawson, E. A., Bui, E., Fischer, L. E., Khadge, S. D., et al. (2014). Gender moderates the effect of oxytocin on social judgments. *Human Psychopharmacology: Clinical and Experimental*, 29(3), 299–304.
- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2), 825–841.
- King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., & Montague, P. R. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science*, 308(5718), 78–83.
- Knafo, A., Israel, S., Darvasi, A., Bachner-Melman, R., Uzefovsky, F., Cohen, L., et al. (2008). Individual differences in allocation of funds in the dictator game associated with length of the arginine vasopressin 1a receptor RS3 promoter region and correlation between RS3 length and hippocampal mRNA. *Genes, Brain and Behavior*, 7(3), 266–275.
- Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, 435(7042), 673–676.
- Krueger, F., Grafman, J., & McCabe, K. (2008). Neural correlates of economic game playing. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 363(1511), 3859–3874.
- Labuschagne, I., Phan, K. L., Wood, A., Angstadt, M., Chua, P., Heinrichs, M., et al. (2010). Oxytocin attenuates amygdala reactivity to fear in generalized social anxiety disorder. *Neuropsychopharmacology*, 35(12), 2403–2413.
- Liu, Y., & Wang, Z. (2003). Nucleus accumbens oxytocin and dopamine interact to regulate pair bond formation in female prairie voles. *Neuroscience*, 121(3), 537–544.
- Loup, F., Tribollet, E., Dubois-Dauphin, M., & Dreifuss, J. (1991). Localization of high-affinity binding sites for oxytocin and vasopressin in the human brain. An autoradiographic study. *Brain Research*, 555(2), 220–232.
- Lynn, S. K., Hoge, E. A., Fischer, L. E., Barrett, L. F., & Simon, N. M. (2014). Gender differences in oxytocin-associated disruption of decision bias during emotion perception. *Psychiatry Research* (ahead of print), 1–22.
- Macdonald, K., & Feifel, D. (2013). Helping oxytocin deliver: considerations in the development of oxytocin-based therapeutics for brain disorders. *Frontiers in Neuroscience*, 7(35), 1–21.
- Marsh, A. A., Henry, H. Y., Pine, D. S., & Blair, R. (2010). Oxytocin improves specific recognition of positive facial expressions. *Psychopharmacology*, 209(3), 225–232.
- Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function*, 214(5–6), 655–667.
- Meyer-Lindenberg, A., Domes, G., Kirsch, P., & Heinrichs, M. (2011). Oxytocin and vasopressin in the human brain: social neuropeptides for translational medicine. *Nature Reviews Neuroscience*, 12(9), 524–538.
- Neumann, I. D., Maloumy, R., Beiderbeck, D. I., Lukas, M., & Landgraf, R. (2013). Increased brain and plasma oxytocin after nasal and peripheral administration in rats and mice. *Psychoneuroendocrinology*, 38(10), 1985–1993.
- Preckel, K., Scheele, D., Kendrick, K. M., Maier, W., & Hurlmann, R. (2014). Oxytocin facilitates social approach behavior in women. *Frontiers in Behavioral Neuroscience*, 8(191), 1–9.
- Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual Review of Psychology*, 62, 23–48.
- Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., & Kilts, C. D. (2002). A neural basis for social cooperation. *Neuron*, 35(2), 395–405.
- Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *NeuroImage*, 22(4), 1694–1703.
- Rilling, J. K., DeMarco, A. C., Hackett, P. D., Thompson, R., Ditzen, B., Patel, R., et al. (2012). Effects of intranasal oxytocin and vasopressin on cooperative behavior and associated brain activity in men. *Psychoneuroendocrinology*, 37(4), 447–461.
- Rilling, J. K., DeMarco, A. C., Hackett, P. D., Chen, X., Gautam, P., Stair, S., et al. (2014). Sex differences in the neural and behavioral response to intranasal oxytocin and vasopressin during human social interaction. *Psychoneuroendocrinology*, 39, 237–248.
- Ross, H. E., & Young, L. J. (2009). Oxytocin and the neural mechanisms regulating social cognition and affiliative behavior. *Frontiers in Neuroendocrinology*, 30(4), 534–547.
- Shahrestani, S., Kemp, A. H., & Guastella, A. J. (2013). The impact of a single administration of intranasal oxytocin on the recognition of basic emotions in humans: a meta-analysis. *Neuropsychopharmacology*, 38(10), 1929–1936.
- Skuse, D. H., & Gallagher, L. (2009). Dopaminergic-neuropeptide interactions in the social brain. *Trends in Cognitive Science*, 13(1), 27–35.
- Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, 17(3), 143–155.
- Sørensen, P., Gjerris, A., & Hammer, M. (1985). Cerebrospinal fluid vasopressin in neurological and psychiatric disorders. *Journal of Neurology, Neurosurgery & Psychiatry*, 48(1), 50–57.
- Striepens, N., Kendrick, K. M., Maier, W., & Hurlmann, R. (2011). Prosocial effects of oxytocin and clinical evidence for its therapeutic potential. *Frontiers in Neuroendocrinology*, 32(4), 426–450.
- Striepens, N., Kendrick, K. M., Hanking, V., Landgraf, R., Wüllner, U., Maier, W., et al. (2013). Elevated cerebrospinal fluid and blood concentrations of oxytocin following its intranasal administration in humans. *Scientific Reports*, 3(3440), 1–5.
- Thompson, R., George, K., Walton, J., Orr, S., & Benson, J. (2006). Sex-specific influences of vasopressin on human social communication. *Proceedings of the National Academy of Sciences*, 103(20), 7889–7894.

- Veenema, A. H., Bredewold, R., & De Vries, G. J. (2013). Sex-specific modulation of juvenile social play by vasopressin. *Psychoneuroendocrinology*, *38*(11), 2554–2561.
- Watanabe, T., Abe, O., Kuwabara, H., Yahata, N., Takano, Y., Iwashiro, N., et al. (2014). Mitigation of sociocommunicational deficits of autism through oxytocin-induced recovery of medial prefrontal activity: a randomized trial. *JAMA Psychiatry*, *71*(2), 166–175.
- Woo, C.-W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. *NeuroImage*, *91*, 412–419.
- Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage*, *14*(6), 1370–1386.
- Young, L. J., Murphy Young, A. Z., & Hammock, E. A. (2005). Anatomy and neurochemistry of the pair bond. *Journal of Comparative Neurology*, *493*(1), 51–57.
- Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Chappelow, J. C., & Berns, G. S. (2004). Human striatal responses to monetary reward depend on saliency. *Neuron*, *42*(3), 509–517.