Null results of oxytocin and vasopressin administration on mentalizing in a large fMRI sample: evidence from a randomized controlled trial

Mark A. Straccia[1], Adam R. Teed[2], Perri L. Katzman[3], Kevin M. Tan[1], Michael H. Parrish[1], Michael R. Irwin[1,4,5], Naomi I. Eisenberger[1], Matthew D. Lieberman[1,4] and Benjamin A. Tabak[2] (ID)

[1]Department of Psychology, University of California, Los Angeles, CA, USA; [2]Department of Psychology, Southern Methodist University, Dallas, TX, USA; [3]Department of Psychology, New York University, New York, NY, USA; [4]Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine, University of California, Los Angeles, CA, USA and [5]Cousins Center for Psychoneuroimmunology, Jane and Terry Semel Institute for Neuroscience, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

Original Article

Abstract

Background. Although potential links between oxytocin (OT), vasopressin (AVP), and social cognition are well-grounded theoretically, most studies have included all male samples, and few have demonstrated consistent effects of either neuropeptide on mentalizing (i.e. understanding the mental states of others). To understand the potential of either neuropeptide as a pharmacological treatment for individuals with impairments in social cognition, it is important to demonstrate the beneficial effects of OT and AVP on mentalizing in healthy individuals.

Methods. In the present randomized, double-blind, placebo-controlled study ($n = 186$) of healthy individuals, we examined the effects of OT and AVP administration on behavioral responses and neural activity in response to a mentalizing task.

Results. Relative to placebo, neither drug showed an effect on task reaction time or accuracy, nor on whole-brain neural activation or functional connectivity observed within brain networks associated with mentalizing. Exploratory analyses included several variables previously shown to moderate OT's effects on social processes (e.g., self-reported empathy, alexithymia) but resulted in no significant interaction effects.

Conclusions. Results add to a growing literature demonstrating that intranasal administration of OT and AVP may have a more limited effect on social cognition, at both the behavioral and neural level, than initially assumed. Randomized controlled trial registrations: ClinicalTrials.gov; NCT02393443; NCT02393456; NCT02394054.

CAMBRIDGE
UNIVERSITY PRESS

Background

A number of oxytocin (OT) and vasopressin (AVP) administration studies point to a possible role for these neuropeptides in improving mentalizing (Brunnlieb, Münte, Tempelmann, & Heldmann, 2013; Domes, Heinrichs, Michel, Berger, & Herpertz, 2007; Tomova, Heinrichs, & Lamm, 2019), the ability to infer and represent others' mental states (Frith & Frith, 1999). Based on results such as these, one of the primary goals of research in this area has been to use OT or AVP as a treatment for individuals with social cognitive impairments (Quintana et al., 2021). However, it is currently unclear whether OT or AVP specifically influences mentalizing. There has been a lack of replication for the behavioral effects of OT and AVP on a range of social processes (Tabak et al., 2019), including mentalizing (Radke & de Bruijn, 2015). In addition, most studies have included all-male samples despite numerous sex-specific effects (Quintana et al., 2021), and surprisingly few have investigated the effects of OT or AVP specifically on mentalizing.

The majority of these studies include the Reading the Mind in the Eyes Test (RMET; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), in which participants identify emotional states from only the eye region of face stimuli. A recent meta-analysis of OT administration studies, many of which included the RMET, found that OT did not affect theory of mind in healthy or clinical samples (Leppanen, Ng, Tchanturia, & Treasure, 2017). However, several studies using the RMET have demonstrated moderation of OT effects by person-specific factors (Feeser et al., 2015; Luminet, Grynberg, Ruzette, & Mikolajczak, 2011; Riem, Bakermans-Kranenburg, Voorthuis, & van IJzendoorn, 2014; Schwaiger, Heinrichs, & Kumsta, 2019; Sun, Vuillier, Deakin, & Kogan, 2020). Importantly, although

CrossMark

the RMET is typically conceptualized as a mentalizing task, there is some ambiguity in this classification as others have viewed it as an emotion recognition task (Adolphs, Baron-Cohen, & Tranel, 2002; Uzefovsky, Shalev, Israel, Knafo, & Ebstein, 2012), and deficits in RMET performance have been disassociated from theory of mind performance (Oakley, Brewer, Bird, & Catmur, 2016). Emotion recognition has been viewed as a process distinct from, overlapping with, or even subsumed by mentalizing, which shares a nearly identical definition with theory of mind, and is typically associated with cognitive rather than emotional mental state inferences (Mitchell & Phillips, 2015).

A few studies have examined the effects of OT on other mentalizing tasks. In healthy control groups, one study ($n = 25$) found no effects of OT in the False-Belief Task, which involves mentalizing in the context of a hypothetical scenario (De Coster, Lin, Mathalon, & Woolley, 2019), and another ($n = 31$) found no effect of OT using The Awareness of Social Inference Test (Woolley et al., 2014), which elicits social inferences from video clips of individuals in contexts of varying complexity (McDonald et al., 2006). Two studies of healthy individuals ($ns = 48–59$) used the Multifaceted Empathy Test (Dziobek et al., 2008) to distinguish between mentalizing, or cognitive empathy (i.e. explicit recognition of others' thoughts and emotion), and affective empathy (i.e. sharing another's perceived emotional state) and found that OT increased affective empathy, but not mentalizing (Geng et al., 2018; Hurlemann et al., 2010). However, Tomova et al. (2019) found OT enhanced mentalizing by sharpening self-other distinction during a perspective taking task.

Although there is limited evidence for a behavioral effect of OT on mentalizing, it is possible that OT affects neural regions associated with mentalizing even in the absence of behavioral changes. The mentalizing network includes the temporoparietal junction (TPJ)/posterior superior temporal sulcus, the medial prefrontal cortex (mPFC), the inferior frontal gyrus (IFG), the temporal poles, and the posterior cingulate cortex (PCC)/precuneus (Mar, 2011; Molenberghs, Johnson, Henry, & Mattingley, 2016; Schurz, Radua, Aichhorn, Richlan, & Perner, 2014; Van Overwalle & Baetens, 2009). To our knowledge, only three studies have used functional magnetic resonance imaging (fMRI) to examine the effect of OT on brain networks that underlie mentalizing in healthy participants. One study found no neural effects of OT on mentalizing using the Multifaceted Empathy Test (Dziobek et al., 2008) in a sample of men and women ($n = 69$) (Geng et al., 2018). However, the only region of the mentalizing network included as a region of interest (ROI) in this study was the mPFC. Another small study ($n = 9$) found some evidence for neural effects of OT on mentalizing using the RMET, but these results were not maintained after contrasting them with a gender identification control task (Pincus et al., 2010). Riem et al. (2014) also used the RMET and found OT increased activation in the insula, IFG superior temporal gyrus, and left paracingulate gyrus using ROI analyses in an all-female sample ($n = 50$). Considering that these three studies either performed ROI analysis that included very few mentalizing network regions, and/or employed the RMET, it remains unclear how OT affects neural activation specifically during mentalizing.

Compared to OT, even less research has examined the effects of AVP on mentalizing. In one study, AVP reduced mentalizing ability in the RMET ($n = 39$ males) (Uzefovsky et al., 2012), and an fMRI study ($n = 39$ males) on inferring emotions during social scenes (i.e. mentalizing) found that AVP increased activity in the PCC, lateral PFC, precentral gyrus, and left insula (Brunnlieb

et al., 2013). These associations align with meta-analyses and reviews that have characterized a network of mentalizing neural correlates (Mar, 2011; Molenberghs et al., 2016; Schurz et al., 2014; Van Overwalle & Baetens, 2009).

In the present study, we recruited a large sample ($n = 186$) to examine the effects of OT or AVP compared to placebo on behavioral responses, neural activation, and functional connectivity when engaged in mentalizing. Participants completed the Why/How Task, which activates a network of mentalizing-related neural regions including the mPFC, lateral orbitofrontal cortex, PCC/precuneus, temporal poles, and the left TPJ (Spunt & Adolphs, 2014; Spunt & Lieberman, 2012b; Spunt, Satpute, & Lieberman, 2011). We also conducted several exploratory analyses including relevant moderators. Although evidence to date suggests that intranasal OT (and possibly AVP) administration may not affect mentalizing, the limitations of the RMET as an assessment of mentalizing point to the possibility that the use of a task that more robustly activates the mentalizing network may elucidate previously unknown effects. In addition, several studies have found effects of OT and AVP when completing tasks that engage social cognition, even if they are not specifically isolating the mentalizing process (Brunnlieb et al., 2016; Declerck, Boone, & Kiyonari, 2010, 2013; Feng et al., 2015; Gozzi, Dashow, Thurm, Swedo, & Zink, 2017; Rilling et al., 2012, 2014; Teed, Han, Rakic, Mark, & Krawczyk, 2019). We therefore hypothesized OT and AVP would modulate activation in, and/or connectivity between, neural regions in the mentalizing network.

Methods

Participants

All participants ($n = 197$) were recruited from the University of California, Los Angeles (UCLA) between January 2015 and July 2016 (see online Supplementary Information for exclusion criteria and online Supplementary Fig. S1 for CONSORT flow chart). The initial sample size collected for the OT ($n = 75$) and placebo ($n = 89$) groups provided 80% power to detect an effect size of 0.45 and 40% power to detect an effect size of 0.27. The initial sample size for AVP provided 80% power to detect an effect size of 0.57 and 40% power to detect an effect size of 0.35.[†1] Five participants did not complete the Why/How Task due to lack of time, or technical issues. Further, participants were removed prior to analysis due to brain abnormalities ($n = 2$), brain alignment issues ($n = 2$), falling asleep in the scanner ($n = 1$), or being the pilot participant ($n = 1$). This resulted in 186 participants (ages 18–28, mean = 20.3, S.D. = 1.72, 119 female) who were randomly assigned in a double-blind procedure to receive OT ($n = 71$; 38 females), AVP ($n = 31$; 31 females), or placebo ($n = 84$; 50 females). Participants identified as Asian (41.4%), White (32.3%), Black or African American (2.2%), Native American or Alaska Native (0.5%), and 'Other' (23.7%). Across all participants, 27.4% identified as Hispanic or Latino. Participants who completed the study were paid $100. Informed consent was obtained from all participants as approved by the UCLA Institutional Review Board. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

[†]The notes appear after the main text.

Fig. 1. Description of Why/How Task. Following a fixation cross presented for an average of 9 s, each of the 16 experimental blocks began with a prompt question shown for 2.1 s followed by a blank screen lasting 0.15 s before presenting a sequence of eight trial images. Participants were given a max of 2.2 s to respond to each image, and a reminder prompt lasting 0.3 s was shown between each image. Each pre-block prompt began with 'Is the person' followed by a descriptive phrase specific to each question. This same phrase was then shown as a reminder between each trial. Recorded BOLD signal was analyzed in a variable epoch manner beginning from the onset of the first image to the offset of the final image of the block. The Why prompts described either intentions inferred from a person's bodily actions (e.g. 'helping someone') or emotional states inferred from facial expressions (e.g. 'proud of themselves'). The How prompts described the physical mechanics of body actions (e.g. 'lifting something') or facial orientations (e.g. 'opening their mouth'). The same set of images were used for both Why and How trials, and participants responded 'yes' or 'no' with their index or middle finger to indicate whether the person(s) in each image demonstrated that mental state or were performing the action stated in the prompt. On average, the task lasted approximately 4.9 min per person. See online Supplemental Information for additional prompt examples and task information.

Procedure

Participants completed two sessions (see online Supplementary Fig. S2 for study timeline). In the first session, participants completed a series of self-report questionnaires. The measures relevant to the present study are described below. The second session occurred 7–145 days later (mean = 45.37, S.D. = 26.41). Random assignment, dosage (i.e. 24 IU/ml for OT and 20 IU/ml for AVP), method of administration, and incubation period were conducted in the same manner as our previous research (Tabak et al., 2015, 2016, 2019; see online Supplemental Information for additional information). Following approximately 40 min of incubation, participants began preparation for the fMRI session. During fMRI scanning, participants performed other tasks (not analyzed for this paper) before the mentalizing task, which occurred approximately 90 min post-administration (as in Bartz et al., 2010).

Why/How task

The Why/How task (Spunt & Lieberman, 2012a, b; Spunt, Meyer, & Lieberman, 2015) reliably activates the mentalizing (Why > How) and mirror (How > Why) networks in the brain. We used a version that was nearly identical to study 3 in Spunt and Adolphs (2014) with minor adjustments to the trial structure timing (see Fig. 1 for additional information). Participants answered *why* (e.g. 'Is the person helping someone') someone is performing an action or *how* (e.g. 'Is the person reaching') someone is performing an action.

Self-report measures

As shown in online Supplementary Table S1, participants completed several self-report measures for exploratory moderation analyses of OT and AVP effects.

fMRI image acquisition

We collected data on a Prisma 3-T MR system at the UCLA Ahmanson-Lovelace Brain Mapping Center. We collected 148 functional volumes using a T2* weighted gradient echo-planar sequence with the following parameters: matrix size = 64×64, $3.1 \times 3.1 \times 3$ mm voxels, repetition time (TR) = 2.0 s, echo time (TE) = 24 ms, flip angle (FA) = 90°, FOV = 1200 mm, bandwidth = 2605 Hz/Px, 20-channel head coil, and no acceleration. Volumes consisted of 36.3 mm slices with a distance factor 33%, and slice orientation tilt of 22.5% relative to the AC/PC plane. At the end of the scanning session, a high-resolution structural volume (MPRAGE) was collected with the sequence parameters: $1.1 \times 1.1 \times 1.2$ mm voxels, TR = 2.3 s, TE = 2.95 ms, FA = 9°, distance factor 50%, and parallel imaging implementation mode GRAPPA with an acceleration factor of 2. The first 96 participants were run with scans going from posterior to anterior. Since dropout of signal can depend on the direction of the scans, for optimal coverage of the ventral PFC, we ran the last 90 participants with scans going from anterior to posterior. Contrasting blood oxygen level-dependent (BOLD) activation for participants whose images had posterior to anterior v. anterior to posterior encoding revealed a signal difference only in the inferior ventromedial prefrontal cortex (vmPFC) signal located behind the frontal sinuses. Thus, we did not expect the coding direction to affect treatment group contrasts elsewhere in the brain.

Statistical analysis

Behavioral analysis

We examined the effects of either neuropeptide on reaction time and accuracy for the Why/How task using two-sided *t* tests and report the 95% confidence intervals (CIs) as well as estimated effect sizes calculated by Cohen's *d*. Accuracy was based on Spunt and Adolphs's (2014) coding of each image. We chose to focus specifically on OT v. placebo or AVP v. placebo because we did not have *a priori* hypotheses about OT v. AVP.

fMRI analysis

General linear model. The imaging data were analyzed in SPM12 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK). Functional volumes were motion and distortion corrected, normalized to a standardized (MNI)
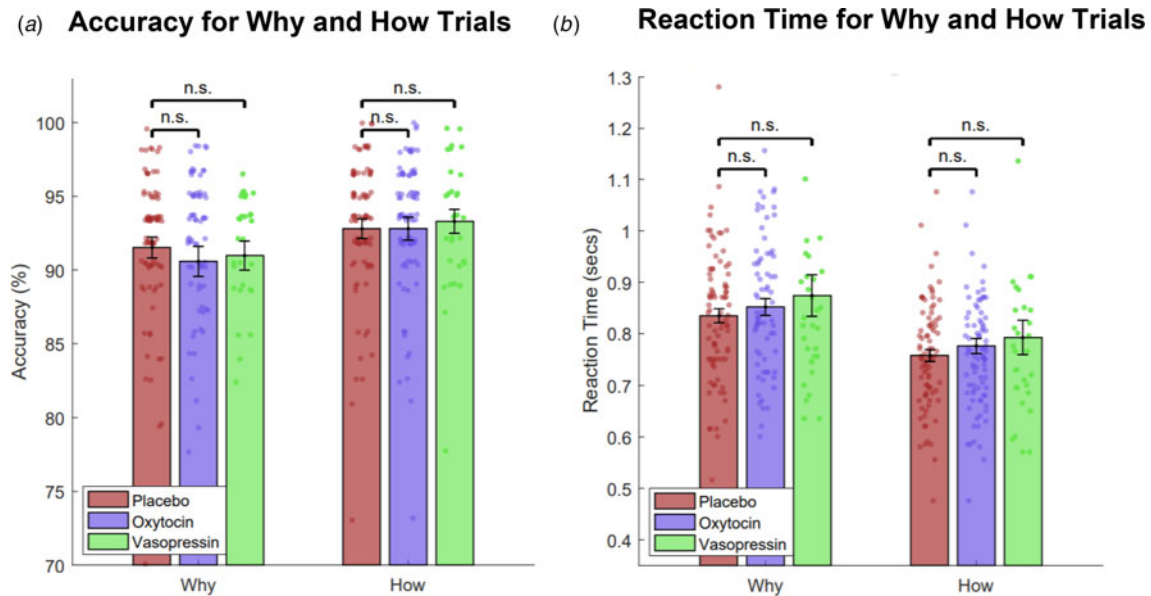
(a) **Accuracy for Why and How Trials**   (b) **Reaction Time for Why and How Trials**

Fig. 2. Differences between OT, AVP, and placebo on accuracy and reaction time. (a) Accuracy for Why and How trials, (b) reaction time for Why and How trials. No significant behavioral differences in accuracy or reaction time were observed for OT and AVP *v.* placebo in either the Why or How trials. Error bars represent standard error. $p < 0.001$***, $p < 0.01$**, $p < 0.05$*, $p < 0.10$†, N.S. = not significant.

template using the DARTEL toolbox (including resampling to 2 mm isotropic voxels), spatially smoothed with a Gaussian kernel (8 mm FWHM) and high-pass filtered (128 s cut-off). Analyses were run on the whole brain as well as on ROIs defined in Spunt and Adolphs (2014). Drug condition (OT, AVP, or placebo) was added as a covariate in the second-level analysis.

GLM Fast Flex 2. To ensure our ability to detect effects in areas prone to drop out, we used GLM Fast Flex 2 (Harvard University, Boston, USA) to conduct whole-brain, voxel-level analysis. This package has the benefit of calculating voxels in which some participants might be missing data. GLM Fast Flex 2 was also used to conduct exploratory moderation analyses involving self-report measures. These self-reports were included as moderators in the second-level analysis.

Functional connectivity. We analyzed the functional connectivity during each task block and during rest using the CONN toolbox (www.nitrc.org/projects/conn; version 17.f). The data were processed with band-pass filtering between 0.01 and 0.1 Hz and the default component-based CompCor method for reducing physiological and other sources of noise (Whitfield-Gabrieli & Nieto-Castanon, 2012).

ROI analyses. Twenty ROIs (10 from the mentalizing network, 10 from the mirror network) were collected from clusters found to be significant at $p < 10^{-6}$ uncorrected from a previously validated version of the Why/How contrast by Spunt and Adolphs (2014). We extracted the β values within each ROI for each participant. Each ROI was tested separately in contrasts for each of the OT or AVP *v.* placebo contrasts. Additionally, to increase signal to noise, *a priori* clusters associated with the mentalizing or mirror networks were grouped and contrasted for group effects as two network ROIs. For the functional connectivity analysis, bivariate correlations from the CONN toolbox were examined between each of the 10 ROIs within the mentalizing and mirror networks respectively. Mentalizing ROIs included the PCC/precuneus, mPFC, vmPFC, left posterior middle frontal gyrus, left and right TPJ, left and right STS, and left and right ventral lateral

PFC. Mirror network ROIs included bilateral intraparietal lobule, bilateral posterior premotor cortex, bilateral posterior IFG, bilateral posterior middle temporal gyrus, and bilateral superior parietal lobule.

All statistical tests were performed in MATLAB (MathWorks, Natick, MA, USA) and report *p* values, CIs, effect size, Bayes factors (BF), and equivalence tests. BFs were calculated as ratios of the marginal likelihoods of the data under the alternative and null hypothesis given a prior odds distribution of $r = 0.707$, the default prior distribution in the BayesFactor package in R from which we extracted code (Morey & Roulder, 2015). The lack of support in the literature for hypothesizing specific behavioral results in this task, or of commonly reported effect sizes in BOLD ROI analyses, limited us to the use of a default prior distribution which favors the null hypothesis in cases where there is a true small effect (Rouder, Speckman, Sun, Morey, & Iverson, 2009). Therefore, the BF results should be viewed with a degree of skepticism.

Results

Oxytocin and vasopressin effects on accuracy and reaction time

We first examined the effect of either neuropeptide on accuracy and reaction time during Why trials (mentalizing) and How trials (action understanding). No significant differences between treatment groups were found in accuracy (Fig. 2a) or reaction time (Fig. 2b) for either mentalizing or action understanding, when comparing either OT or AVP to placebo (see online Supplemental results).

Oxytocin and vasopressin effects on neural activation

We then examined the effects of either neuropeptide on brain activation during the Why/How task as measured by BOLD

(a) Neural Activation of Why vs How
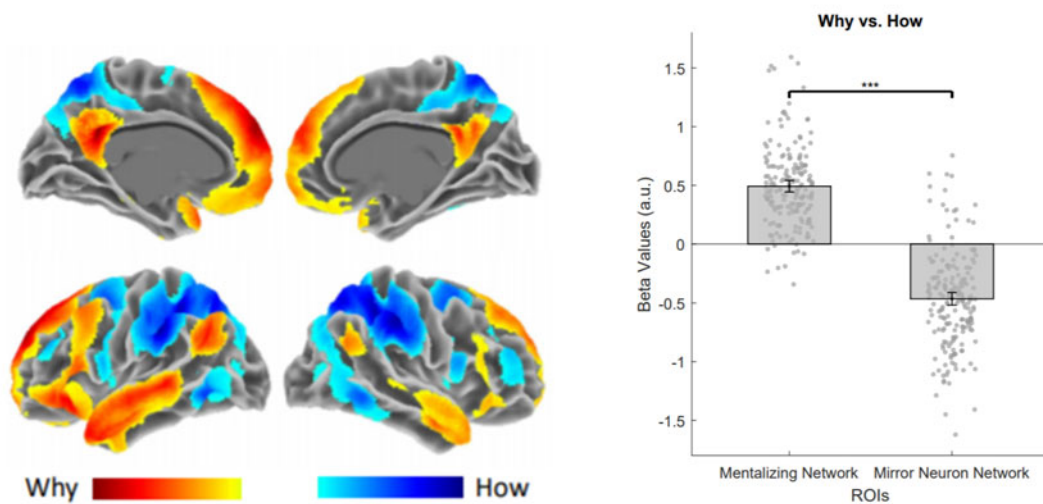
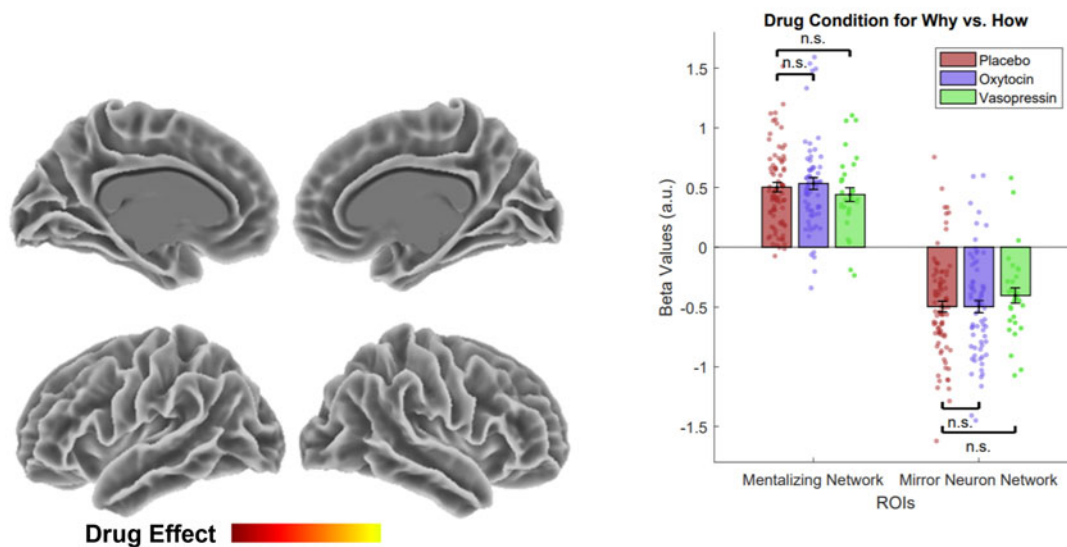## (b) Neural Activation of Drug Condition for Why vs How

Fig. 3. Differences in neural activity found via *t* tests contrasting the effects of either OT or AVP *v.* placebo. (a) Neural activation of Why *v.* How, (b) neural activation of drug condition for Why *v.* How. (a) The contrast of Why *v.* How activation collapsed across all conditions shows robust activation in the mentalizing network ROI for Why trials and mirror network ROI for How trials. Warmer colors correspond to greater activation during Why trials and cooler colors to greater activation during How trials. (b) The graph on the right shows there is no significant difference in neural activation for either OT or AVP *v.* placebo ($p > 0.05$) in the mentalizing network ROI for Why trials and mirror network ROI for How trials. The maps on the left show no significant activation ($p < 0.001$ uncorrected) for either drug condition. Error bars represent standard error. $p < 0.001$***, $p < 0.01$**, $p < 0.05$*, $p < 0.10$†, N.S. = not significant.

response. We first analyzed the main effect of the Why *v.* How contrast across all participants (including OT, AVP, and placebo groups). Replicating prior work, we found robust activation of the mentalizing network ROI for the Why compared to How trials in the placebo group [$t_{83} = 12.621$, $p < 1 \times 10^{-10}$, 95% CI (0.42–0.58), Cohen's $d = 1.377$] and when collapsing across all drug conditions [$t_{185} = 18.1$, $p < 1 \times 10^{-10}$, 95% CI (0.45–0.56), Cohen's $d = 1.33$] (Fig. 3a). We also found robust activation of the mirror network ROI for How compared to Why trials in the placebo group [$t_{83} = 10.82$, $p < 1 \times 10^{-10}$, 95% CI (0.59–0.40), Cohen's $d = 1.181$] and when collapsing across all drug conditions [$t_{185} = 15.893$, $p < 1 \times 10^{-10}$, 95% CI (0.54–0.42), Cohen's $d = 1.165$] (Fig. 3a).

We next examined the effects of OT or AVP *v.* placebo on the Why *v.* How trials. When conducting whole-brain analyses, no effect was found for OT or AVP compared to placebo for the whole-brain contrast at a liberal threshold of $p < 0.001$ uncorrected at either the voxel or cluster level (Fig. 3b). Next, when comparing β values for the mentalizing *v.* mirror networks, we found no effect of Why *v.* How in the mentalizing network for OT *v.* placebo or AVP *v.* placebo, nor did we find any effect for the reverse How *v.* Why contrast in the mirror network for OT *v.* placebo or AVP *v.* placebo (Fig. 3b). Finally, when examining each of the ROIs within the networks, we also found no effects of OT *v.* placebo (Table 1) or AVP *v.* placebo (Table 2) in any ROI of either the mentalizing or mirror networks (see online

Table 1. Neural region of interest signal change differences between oxytocin and placebo groups as tested via both *t* tests and Bayes factors

ROI name	Null hypothesis significance testing							Bayes factor	
	df	*t*-value	*p* value uncorrected	*p* value FDR corrected	Lower 95% CI	Upper 95% CI	Cohen's *d*	BF01	BF10
Why network	139.6	0.482	0.630	0.990	−0.096	0.157	0.079	5.17	0.193
PCC	148.8	0.304	0.761	0.980	−0.108	0.147	0.049	5.52	0.181
mPFC	150.1	0.038	0.970	0.980	−0.114	0.119	0.006	5.76	0.174
vmPFC	145.3	0.062	0.951	0.980	−0.095	0.101	0.010	5.75	0.174
LpMFG	151.0	0.315	0.753	0.980	−0.118	0.163	0.051	5.50	0.182
LTPJ	138.4	0.234	0.815	0.980	−0.150	0.190	0.038	5.62	0.178
RTPJ	151.4	0.036	0.971	0.980	−0.201	0.209	0.006	5.76	0.174
LSTS	150.0	1.074	0.285	0.980	−0.031	0.105	0.173	3.39	0.295
RSTS	142.6	0.050	0.960	0.980	−0.062	0.066	0.008	5.75	0.174
LvlPFC	147.5	0.421	0.675	0.980	−0.071	0.110	0.068	5.31	0.188
RvlPFC	139.7	0.026	0.980	0.980	−0.084	0.086	0.004	5.76	0.174
How network	147.7	0.013	0.990	0.990	−0.134	0.136	0.002	5.76	0.174
LIPL	144.4	0.346	0.730	0.980	−0.105	0.149	0.056	5.45	0.183
RIPL	143.4	0.996	0.321	0.980	−0.060	0.181	0.162	3.65	0.274
LpPMC	134.5	0.475	0.635	0.980	−0.100	0.164	0.078	5.19	0.193
RpPMC	142.2	0.968	0.335	0.980	−0.069	0.202	0.157	3.74	0.267
LpIFG	131.1	0.148	0.883	0.980	−0.099	0.115	0.024	5.70	0.175
RpIFG	147.2	0.196	0.845	0.980	−0.084	0.102	0.032	5.66	0.177
LpMTG	142.0	0.271	0.787	0.980	−0.137	0.181	0.044	5.57	0.180
RpMTG	146.9	0.673	0.502	0.980	−0.117	0.238	0.109	4.68	0.214
LSPL	148.2	0.325	0.745	0.980	−0.162	0.226	0.052	5.49	0.182
RSPL	148.7	0.768	0.444	0.980	−0.121	0.274	0.124	4.39	0.228
Amygdala	150.9	0.359	0.720	0.990	−0.073	0.105	0.056	5.43	0.184

Supplementary Fig. S3 for equivalence test results). AVP did demonstrate a positive moderate effect on each of the left and right premotor cortices of the mirror network compared to placebo when How was compared to Why trials (Left: BF10 = 4.92; Right: BF10 = 7.39), but this did not maintain significance following multiple test correction using the False Discovery Rate (FDR; Benjamini & Hochberg, 1995) in *t* tests [Left: $t_{56} = 2.700$, $p_{adj} = 0.09$, 95% CI (0.064–0.432), Cohen's $d = 0.642$; Right: $t_{65} = 2.872$, $p_{adj} = 0.09$, 95% CI (0.075–0.418), Cohen's d = 0.652].

Oxytocin and vasopressin effects on functional connectivity

We also examined functional connectivity while the participants were performing the task. First, we measured how changes in connectivity for the Why *v.* How contrast differed for each of the OT and AVP conditions compared to the placebo group. The connectivity between 20 ROIs provided 190 comparisons for both drugs. After FDR correction, we found no significant difference in connectivity during Why compared to How trials for either OT or AVP *v.* placebo between any of the 20 ROIs (Fig. 4*a* and *b*; online Supplementary Fig. S3).

Exploratory analyses of gender

We further examined potential gender interaction effects with OT *v.* placebo on accuracy and reaction time for Why/Why and How/How trials and found no effects (see online Supplementary Table S2 and online Supplementary Figs S3 and S4). We also found no significant differences in whole-brain and ROI analyses, nor functional connectivity in men or women (online Supplementary Figs S3, S5, and S6).

Exploratory moderation analyses

To explore potentially relevant moderators of OT or AVP effects, we separately averaged the activity observed within each of the mentalizing and mirror networks from Spunt and Adolphs (2014). To aid interpretation, we *z*-scored the neural data and each moderator across participants and included each moderator individually in a regression as an interaction with the dummy-coded drug conditions. We examined comparisons using both a liberal nominal significance (*p* < 0.01 uncorrected) and the FDR correction (*q* = 0.05). No interaction effects were observed for the mentalizing network. We found an interaction only for the Autism Spectrum

Table 2. Neural region of interest signal change differences between vasopressin and placebo groups as tested via both *t* tests and Bayes factors

ROI name	df	*t*-value	*p* value uncorrected	*p* value FDR corrected	Lower 95% CI	Upper 95% CI	Cohen's *d*	BF01	BF10
			Null hypothesis significance testing					Bayes factor	
Why network	57.1	0.459	0.648	0.648	−0.109	0.173	0.108	3.85	0.259
PCC	59.0	1.019	0.312	0.481	−0.101	0.311	0.238	2.70	0.371
mPFC	55.8	0.253	0.801	0.890	−0.168	0.217	0.060	4.11	0.243
vmPFC	58.8	0.458	0.649	0.865	−0.122	0.194	0.107	3.86	0.259
LpMFG	61.0	0.499	0.620	0.865	−0.149	0.248	0.116	3.79	0.264
LTPJ	53.5	1.047	0.300	0.481	−0.115	0.367	0.252	2.63	0.381
RTPJ	59.9	1.019	0.312	0.481	−0.155	0.478	0.237	2.70	0.371
LSTS	65.4	0.280	0.780	0.890	−0.090	0.120	0.064	4.08	0.245
RSTS	60.5	0.129	0.898	0.945	−0.083	0.094	0.030	4.19	0.238
LvlPFC	66.3	0.004	0.997	0.997	−0.132	0.132	0.001	4.22	0.237
RvlPFC	56.5	0.353	0.725	0.890	−0.104	0.149	0.084	4.00	0.250
How network	60.4	0.971	0.335	0.648	−0.081	0.234	0.226	2.81	0.356
LIPL	57.4	2.003	0.050	0.184	0.000	0.365	0.472	0.77	1.303
RIPL	50.8	1.698	0.096	0.217	−0.031	0.369	0.415	1.23	0.811
LpPMC	55.7	2.700	**0.009**	0.092	0.064	0.432	0.642	0.20	**4.918**
RpPMC	65.1	2.872	**0.005**	0.092	0.075	0.418	0.652	0.14	**7.393**
LpIFG	62.2	1.281	0.205	0.410	−0.046	0.209	0.295	2.08	0.480
RpIFG	58.3	1.758	0.084	0.217	−0.015	0.234	0.413	1.13	0.887
LpMTG	59.5	1.683	0.098	0.217	−0.038	0.445	0.393	1.26	0.796
RpMTG	57.5	1.956	0.055	0.184	−0.006	0.510	0.461	0.83	1.207
LSPL	57.4	2.342	**0.023**	0.151	0.047	0.596	0.552	0.42	2.388
RSPL	53.4	2.105	**0.040**	0.184	0.015	0.597	0.507	0.65	1.545
Amygdala	61.8	0.613	0.542	0.648	−0.102	0.193	0.141	3.59	0.279

Note. Bold = *p* < 0.05 and/or BF10>3.

Quotient × OT *v.* placebo in the mirror network [$b = -0.494$, $t(180) = -2.99$, $p = 0.0032$] that showed nominal significance ($p < 0.01$ uncorrected) (online Supplementary Fig. S7), but this did not survive multiple test correction (FDR corrected $p = 0.14$).

Discussion

In this large fMRI study, we found no effects of OT or AVP *v.* placebo on either neural activation or functional connectivity when contrasting inferences of Why *v.* How a behavior was performed. We also observed no behavioral differences between OT or AVP *v.* placebo for task accuracy or response times. The absence of effects is noteworthy considering that highly significant differences in neural activation between Why/How conditions were found when collapsing across drug and placebo groups. Thus, these null findings do not appear to result from a failure of the Why/How task to elicit mentalizing or mirror networks. Rather, these findings represent a lack of evidence that either neuropeptide influences mentalizing at the behavioral or neural level.

Meta-analyses have found evidence of beneficial effects of OT on theory of mind in studies of people with neurodevelopmental disorders (Bürkner, Williams, Simmons, & Woolley, 2017; Keech, Crowe, & Hocking, 2018), and preliminary evidence suggests that AVP administration may improve social ability in children with autism spectrum disorders (Parker et al., 2019). In line with these findings, studies with healthy samples have found OT enhanced performance on social tasks among individuals with lower levels of self-reported social cognitive ability (Bartz et al., 2010, 2019; Feeser et al., 2015; Luminet et al., 2011; Radke & de Bruijn, 2015). Given these prior findings, we examined several self-report measures of social cognitive ability, as well as other relevant potential moderators, but found no evidence to support moderation effects following correction for multiple testing.

The lack of results for OT or AVP suggests that these neuropeptides may not influence mentalizing in healthy individuals. This is consistent with the results from Geng et al. (2018) who found null results for mentalizing using the Multifaceted Empathy Task (Dziobek et al., 2008) which features contextually and visually rich stimuli similar to those in the Why/How task presented here. We also extended the findings of Geng et al. (2018) to regions beyond the mPFC, amygdala and anterior insula by conducting a whole-brain analysis as well as targeted

(a) Significant Functional Connectivity Changes in the Why/How Contrast for OT vs. Placebo

(b) Significant Functional Connectivity Changes in the Why/How Contrast for AVP vs. Placebo

p<0.05 (FDR corrected)
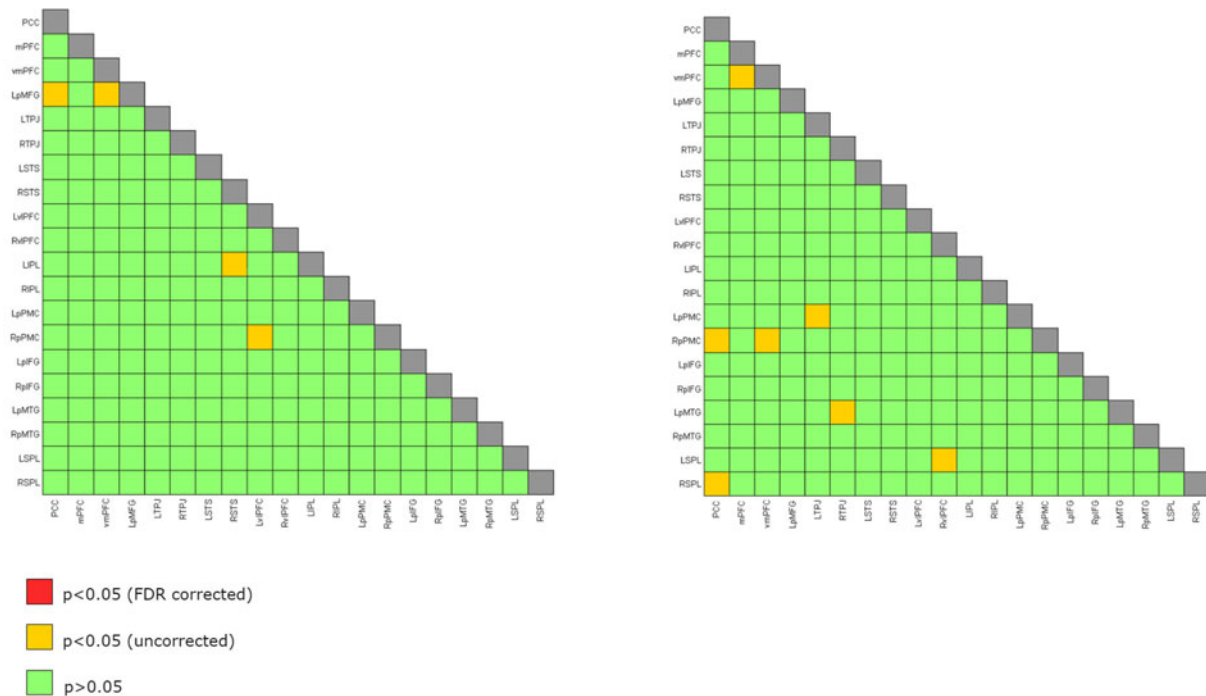
p<0.05 (uncorrected)

p>0.05

Fig. 4. Twenty ROIs chosen from Spunt and Adolphs (2014). (a) Significant functional connectivity changes in the Why/How contrast for OT *v.* placebo. (b) Significant functional connectivity changes in the Why/How contrast tasks for AVP *v.* placebo. *Note.* The ROIs used are very similar to the significant activations shown in Fig. 3a (Why *v.* How contrast). (a) No significant difference in functional connectivity during Why compared to How trials for OT *v.* placebo. (b) No significant difference in functional connectivity during Why compared to How trials for AVP (only females) *v.* placebo (only females). Green squares for *p* > 0.05, yellow squares for *p* < 0.05 uncorrected for multiple comparisons, red squares for *p* < 0.05 FDR corrected, and gray squares for correlations of ROI with itself. PCC, posterior cingulate cortex/precuneus; mPFC, medial prefrontal cortex, vmPFC, ventromedial prefrontal cortex; LpMFG, left posterior middle frontal gyrus; LTPJ, left temporoparietal junction; RTPJ, right temporoparietal junction; RSTS, right superior temporal sulcus; LSTS, left superior temporal sulcus; LvlPFC, left ventral lateral prefrontal cortex; RvlPFC, right ventral lateral prefrontal cortex; LIPL, left intraparietal lobule; RIPL, right intraparietal lobule; LpPMC, left posterior premotor cortex; RpPMC, right posterior premotor cortex; LpIFG, left posterior inferior frontal gyrus; RpIFG, right posterior inferior frontal gyrus; LpMTG, left posterior middle temporal gyrus; RpMTG, right posterior middle temporal gyrus; LSPL, left superior parietal lobule; RSPL, right superior parietal lobule.

analyses for several regions in the mentalizing and mirror networks.

It must be noted that there is still skepticism about the use of intranasal OT and AVP administration due to methodological issues, and in particular, the unclear pharmacokinetics (Leng & Ludwig, 2016; but see Quintana et al., 2021 for recent advances). Furthermore, failed replications (Declerck, Boone, Pauwels, Vogt, & Fehr, 2020; Nave, Camerer, & McCullough, 2015), a publication bias against null results (Lane, Luminet, Nave, & Mikolajczak, 2016; Tabak et al., 2019), and issues of statistical power have promoted uncertainty regarding all but the most consistent findings in the OT literature (Mierop et al., 2020; Walum, Waldman, & Young, 2016). To date, the effects of OT appear more robust for basic processes such as emotion recognition (Leppanen et al., 2017) and attention orientation toward social cues (Eckstein et al., 2019; Guastella, Mitchell, & Dadds, 2008; Hubble et al., 2017). While these processes are certainly relevant for mentalizing, they are not one and the same as engaging in mentalizing. Thus, it is possible that OT plays a greater role in the basic mechanisms that are foundational for social cognitive ability (Mitchell & Phillips, 2015).

The present results may also differ from previous studies using the RMET because it is possible that OT may more strongly affect emotion recognition than mentalizing processes. Although

mentalizing and emotion recognition networks show overlap in the mPFC (Lieberman, Straccia, Meyer, Du, & Tan, 2019) and the IFG (Hooker, Verosky, Germine, Knight, & D'Esposito, 2008), emotion recognition tasks are typically more associated with activity in limbic and paralimbic regions such as the amygdala, anterior insula, and parahippocampal gyrus, as well as the fusiform face area and other visual processing areas (Dricu & Frühholz, 2016; Fusar-Poli et al., 2009; Sabatinelli et al., 2011). Based on meta-analytic evidence showing that OT tends to affect limbic regions, particularly the amygdala and insula (Grace, Rossell, Heinrichs, Kordsachia, & Labuschagne, 2018; Zink & Meyer-Lindenberg, 2012), emotion recognition tasks that engage these regions may be more susceptible to OT's influence, whereas mentalizing tasks that engage broader cortical networks may be less affected.

Another possibility for the present findings is that the main effects of the Why/How task are too strong to see drug effects. Indeed, one of the primary strengths of the Why/How task compared to other mentalizing paradigms, such as the RMET, is the tight control it offers in parsing neural effects related to mental state inference *v.* inferences of the physical mechanics of behaviors signaling those mental states (Spunt, Falk, & Lieberman, 2010). However, the near ceiling behavioral accuracy of the Why/How task (i.e. >90%) is also a limiting factor in our analysis,

particularly in relation to identifying behavioral effects of OT or AVP. As such, future studies examining the behavioral effects of OT or AVP on mentalizing would benefit from the inclusion of a similarly well-validated mentalizing task that is more challenging for healthy individuals.

The present study has a number of strengths including the use of one of the most well-validated fMRI tasks for reliably dissociating the mentalizing and mirror networks, indicating the rare ability amongst fMRI tasks of distinguishing mental state representations from perceptual and motor representations. We also demonstrate null results of both main effects and interaction effects related to measures previously observed to moderate the effects of OT or AVP on social cognitive processes.

Since the majority of human studies examining the effects of OT and AVP has relied on all male samples, our majority female sample can be viewed as a strength and represents an important contribution to the literature (Quintana et al., 2021). At the same time, the inclusion of a majority female sample prevents us from knowing the extent to which these results would be confirmed in a sample with an equal number of male participants. In addition, recruiting only healthy participants does not allow us to know whether the present results may generalize to those with psychiatric disorders. The present results also relied on a single, conventional dose of OT and AVP, and it is unclear whether results would remain consistent with a lower or higher dose of either neuropeptide.

In addition, our results only relate to the effects of OT or AVP approximately 90 min post-administration. The current consensus for the optimal amount of time for incubation is approximately 35–50 min before beginning the task. Therefore, our results may not represent the peak increase in OT concentrations (Quintana et al., 2021). Nonetheless, studies have used similar incubation periods and found effects of OT (Bartz et al., 2010). There is also evidence that OT remains increased in cerebrospinal fluid and blood plasma for 75–85 min and beyond following OT administration (Spengler et al., 2017; Striepens et al., 2013), and intranasal OT has produced sustained changes in resting regional cerebral blood flow in numerous regions implicated in socioemotional processing through 78 min post-administration (Paloyelis et al., 2016). Future studies are needed to examine if the null results found in the present study extend to shorter or longer incubation periods, as well as larger or repeated doses.

Another limitation is that our between-subjects design limited overall power (Quintana et al., 2021). Future studies may consider a within-subjects design to increase power. In addition, although we observed general agreement regarding our null findings across different types of analyses, inconsistencies in the OT administration literature prevented us from using a better-informed prior distribution. Thus, we were limited to using a default prior distribution. The smaller a potential true effect is, the more the default prior distribution can be considered to overweigh larger effect sizes. Small effects are then farther from the average distribution effect and less distinguishable from zero, introducing bias toward the null hypothesis (Rouder et al., 2009). Thus, since the effects of OT and AVP in the Why/How task may be small, and we cannot know how small, the interpretation of BF results should be qualified accordingly.

Last, it is important to note that in our preregistration in 2015, the Why/How task was overlooked and not listed because it was initially intended as a localizer for use in a separate task in the same broader study (a task described as 'learning for teaching'). In addition, unfortunately, for logistical purposes, the larger study was separated into three different preregistrations based on different funding mechanisms (NCT02393443; NCT02393456; NCT02394054). Information on the overall study design and the other tasks, including an empathy task, a deception detection task, and a task involving viewing images of participants' attachment figures, were listed in the original preregistrations. Null results that are not preregistered, such as those in the present study, may be subject to biases such as reverse *p*-hacking (Chuard, Vrtílek, Head, & Jennions, 2019). Nevertheless, the results ran counter to our hypotheses and extensive analyses were completed to allow OT and AVP effects to present themselves.

In sum, in a majority female non-clinical sample, we found no effects of either OT or AVP administration compared to placebo on behavioral responses, neural activation, or functional connectivity related to mentalizing, a social cognitive process that is impaired in several clinical populations. Furthermore, of the relevant moderators we explored, no associations survived after correction for multiple comparisons. These results, based on the well-validated Why/How task, support previous research showing a lack of an association between OT and several social processes (Tabak et al., 2019), including mentalizing, at the behavioral level (Radke & de Bruijn, 2015). Moreover, by examining whole-brain and network effects, we add to prior work that has shown a lack of an association at the neural ROI level.

Notes

[1] Based on time and resources, we had the opportunity to include AVP administration for approximately 30 individuals. Given that we could only include this number, we chose to include only female participants in this condition because the vast majority of AVP administration studies have only included male participants (Brunnlieb et al., 2013; Gozzi et al., 2017; Thijssen et al., 2018).

References

Adolphs, R., Baron-Cohen, S., & Tranel, D. (2002). Impaired recognition of social emotions following amygdala damage. *Journal of Cognitive Neuroscience*, 14(8), 1264–1274. doi: 10.1162/089892902760807258

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, 42(2), 241–251. doi: 10.1111/1469-7610.00715

Bartz, J. A., Nitschke, J. P., Krol, S. A., & Tellier, P. P. (2019). Oxytocin selectively improves empathic accuracy: A replication in men and novel insights in women. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 4(12), 1042–1048. doi: 10.1016/j.bpsc.2019.01.014

Bartz, J. A., Zaki, J., Ochsner, K. N., Bolger, N., Kolevzon, A., Ludwig, N., … Taylor, S. E. (2010). Effects of oxytocin on recollections of maternal care

and closeness. *Proceedings of the National Academy of Sciences of the USA*, *107*(50), 21371–21375. doi: 10.1073/pnas.1012669107

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, *57*(1), 289–300. doi: 10.2307/2346101

Brunnlieb, C., Münte, T. F., Tempelmann, C., & Heldmann, M. (2013). Vasopressin modulates neural responses related to emotional stimuli in the right amygdala. *Brain Research*, *1499*, 29–42. doi: 10.1016/j.brainres.2013.01.009

Brunnlieb, C., Nave, G., Camerer, C. F., Schosser, S., Vogt, B., Münte, T. F., & Heldmann, M. (2016). Vasopressin increases human risky cooperative behavior. *Proceedings of the National Academy of Sciences of the USA*, *113*(8), 2051–2056. doi: 10.1073/pnas.1518825113

Bürkner, P.-C., Williams, D. R., Simmons, T. C., & Woolley, J. D. (2017). Intranasal oxytocin may improve high-level social cognition in schizophrenia, but not social cognition or neurocognition in general: A multilevel Bayesian meta-analysis. *Schizophrenia Bulletin*, *43*(6), 1291–1303. doi: 10.1093/SCHBUL/SBX053

Chuard, P. J. C., Vrtílek, M., Head, M. L., & Jennions, M. D. (2019). Evidence that nonsignificant results are sometimes preferred: Reverse P-hacking or selective reporting?. *PLoS Biology*, *17*(1), e3000127. doi: 10.1371/journal.pbio.3000127

Declerck, C. H., Boone, C., & Kiyonari, T. (2010). Oxytocin and cooperation under conditions of uncertainty: The modulating role of incentives and social information. *Hormones and Behavior*, *57*(3), 368–374. doi: 10.1016/j.yhbeh.2010.01.006

Declerck, C. H., Boone, C., & Kiyonari, T. (2013). The effect of oxytocin on cooperation in a prisoner's dilemma depends on the social context and a person's social value orientation. *Social Cognitive and Affective Neuroscience*, *9*(6), 802–809. doi: 10.1093/scan/nst040

Declerck, C. H., Boone, C., Pauwels, L., Vogt, B., & Fehr, E. (2020). A registered replication study on oxytocin and trust. *Nature Human Behaviour*, *4*(6), 646–655. doi: 10.1038/s41562-020-0878-x

De Coster, L., Lin, L., Mathalon, D. H., & Woolley, J. D. (2019). Neural and behavioral effects of oxytocin administration during theory of mind in schizophrenia and controls: A randomized control trial. *Neuropsychopharmacology*, *44*(11), 1925–1931. doi: 10.1038/s41386-019-0417-5

Domes, G., Heinrichs, M., Michel, A., Berger, C., & Herpertz, S. C. (2007). Oxytocin improves 'mind-reading' in humans. *Biological Psychiatry*, *61*(6), 731–733. doi: 10.1016/j.biopsych.2006.07.015

Dricu, M., & Frühholz, S. (2016). Perceiving emotional expressions in others: Activation likelihood estimation meta-analyses of explicit evaluation, passive perception and incidental perception of emotions. *Neuroscience and Biobehavioral Reviews*, *71*, 810–828. doi: 10.1016/j.neubiorev.2016.10.020

Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H. R., Wolf, O. T., & Convit, A. (2008). Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). *Journal of Autism and Developmental Disorders*, *38*(3), 464–473. doi: 10.1007/s10803-007-0486-x

Eckstein, M., Bamert, V., Stephens, S., Wallen, K., Young, L. J., Ehlert, U., & Ditzen, B. (2019). Oxytocin increases eye-gaze towards novel social and non-social stimuli. HHS public access. *Social Neuroscience*, *14*(5), 594–607. doi: 10.1080/17470919.2018.1542341

Feeser, M., Fan, Y., Weigand, A., Hahn, A., Gärtner, M., Böker, H., … Bajbouj, M. (2015). Oxytocin improves mentalizing – pronounced effects for individuals with attenuated ability to empathize. *Psychoneuroendocrinology*, *53*, 223–232. doi: 10.1016/j.psyneuen.2014.12.015

Feng, C., Hackett, P. D., DeMarco, A. C., Chen, X., Stair, S., Haroon, E., … Rilling, J. K. (2015). Oxytocin and vasopressin effects on the neural response to social cooperation are modulated by sex in humans. *Brain Imaging and Behavior*, *9*(4), 754–764. doi: 10.1007/s11682-014-9333-9

Frith, C. D., & Frith, U. (1999). Interacting minds – A biological basis. *Science (New York, N.Y.)*, *286*(5445), 1692–1695. doi: 10.1126/science.286.5445.1692

Fusar-Poli, P., Placentino, A., Carletti, F., Landi, P., Allen, P., Surguladze, S., … Politi, P. (2009). Functional atlas of emotional faces processing: A voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *Journal of Psychiatry and Neuroscience*, *34*(6), 418–432.

Geng, Y., Zhao, W., Zhou, F., Ma, X., Yao, S., Hurlemann, R., … Kendrick, K. M. (2018). Oxytocin enhancement of emotional empathy: Generalization across cultures and effects on amygdala activity. *Frontiers in Neuroscience*, *12*, 512. doi: 10.3389/fnins.2018.00512

Gozzi, M., Dashow, E. M., Thurm, A., Swedo, S. E., & Zink, C. F. (2017). Effects of oxytocin and vasopressin on preferential brain responses to negative social feedback. *Neuropsychopharmacology*, *42*(7), 1409–1419. doi: 10.1038/npp.2016.248

Grace, S. A., Rossell, S. L., Heinrichs, M., Kordsachia, C., & Labuschagne, I. (2018). Oxytocin and brain activity in humans: A systematic review and coordinate-based meta-analysis of functional MRI studies. *Psychoneuroendocrinology*, *96*, 6–24. doi: 10.1016/j.psyneuen.2018.05.031

Guastella, A. J., Mitchell, P. B., & Dadds, M. R. (2008). Oxytocin increases gaze to the eye region of human faces. *Biological Psychiatry*, *63*(1), 3–5. doi: 10.1016/j.biopsych.2007.06.026

Hooker, C. I., Verosky, S. C., Germine, L. T., Knight, R. T., & D'Esposito, M. (2008). Mentalizing about emotion and its relationship to empathy. *Social Cognitive and Affective Neuroscience*, *3*(3), 204–217. doi: 10.1093/scan/nsn019

Hubble, K., Daughters, K., Manstead, A. S. R., Rees, A., Thapar, A., & van Goozen, S. H. M. (2017). Oxytocin increases attention to the eyes and selectively enhances self-reported affective empathy for fear. *Neuropsychologia*, *106*, 350–357. doi: 10.1016/j.neuropsychologia.2017.10.019

Hurlemann, R., Patin, A., Onur, O. A., Cohen, M. X., Baumgartner, T., Metzler, S., … Kendrick, K. M. (2010). Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. *Journal of Neuroscience*, *30*(14), 4999–5007. doi: 10.1523/JNEUROSCI.5538-09.2010

Keech, B., Crowe, S., & Hocking, D. R. (2018). Intranasal oxytocin, social cognition and neurodevelopmental disorders: A meta-analysis. *Psychoneuroendocrinology*, *87*, 9–19. doi: 10.1016/j.psyneuen.2017.09.022

Lane, A., Luminet, O., Nave, G., & Mikolajczak, M. (2016). Is there a publication bias in behavioural intranasal oxytocin research on humans? Opening the file drawer of one laboratory. *Journal of Neuroendocrinology*, *28*(4), 1–15. doi: 10.1111/jne.12384

Leng, G., & Ludwig, M. (2016). Intranasal oxytocin: Myths and delusions. *Biological Psychiatry*, *79*(3), 243–250. doi: 10.1016/j.biopsych.2015.05.003

Leppanen, J., Ng, K. W., Tchanturia, K., & Treasure, J. (2017). Meta-analysis of the effects of intranasal oxytocin on interpretation and expression of emotions. *Neuroscience & Biobehavioral Reviews*, *78*, 125–144. doi: 10.1016/j.neubiorev.2017.04.010

Lieberman, M. D., Straccia, M. A., Meyer, M. L., Du, M., & Tan, K. M. (2019). Social, self, (situational), and affective processes in medial prefrontal cortex (MPFC): Causal, multivariate, and reverse inference evidence. *Neuroscience and Biobehavioral Reviews*, *99*, 311–328. doi: 10.1016/j.neubiorev.2018.12.021

Luminet, O., Grynberg, D., Ruzette, N., & Mikolajczak, M. (2011). Personality-dependent effects of oxytocin: Greater social benefits for high alexithymia scorers. *Biological Psychology*, *87*(3), 401–406. doi: 10.1016/j.biopsycho.2011.05.005

Mar, R. A. (2011). The neural bases of social cognition and story comprehension. *Annual Review of Psychology*, *62*(1), 103–134. doi: 10.1146/annurev-psych-120709-145406

McDonald, S., Bornhofen, C., Shum, D., Long, E., Saunders, C., & Neulinger, K. (2006). Reliability and validity of The Awareness of Social Inference Test (TASIT): A clinical test of social perception. *Disability and Rehabilitation*, *28*(24), 1529–1542. doi: 10.1080/09638280600646185

Mierop, A., Mikolajczak, M., Stahl, C., Béna, J., Luminet, O., Lane, A., & Corneille, O. (2020). How can intranasal oxytocin research be trusted? A systematic review of the interactive effects of intranasal oxytocin on psychosocial outcomes. *Perspectives on Psychological Science*, *15*(5), 1228–1242. doi: 10.1177/1745691620921525

Mitchell, R. L. C., & Phillips, L. H. (2015). The overlapping relationship between emotion perception and theory of mind. *Neuropsychologia*, *70*, 1–10. doi: 10.1016/J.NEUROPSYCHOLOGIA.2015.02.018

Molenberghs, P., Johnson, H., Henry, J. D., & Mattingley, J. B. (2016). Understanding the minds of others: A neuroimaging meta-analysis. *Neuroscience and Biobehavioral Reviews*, *65*, 276–291, doi:10.1016/j.neubiorev.2016.03.020.

Morey, R. D., & Roulder, J. N. (2015). BayesFactor: Computation of Bayes factors for common designs. Retrieved from http://cran/r-projectorg/web/packages/BayesFactor/.

Nave, G., Camerer, C., & McCullough, M. (2015). Does oxytocin increase trust in humans? A critical review of research. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 10(6), 772–789. doi: 10.1177/1745691615600138

Oakley, B. F. M., Brewer, R., Bird, G., & Catmur, C. (2016). Theory of mind is not theory of emotion: A cautionary note on the reading the mind in the eyes test. *Journal of Abnormal Psychology*, 125(6), 818–823. doi: 10.1037/abn0000182

Paloyelis, Y., Doyle, O. M., Zelaya, F. O., Maltezos, S., Williams, S. C., Fotopoulou, A., & Howard, M. A. (2016). A spatiotemporal profile of in vivo cerebral blood flow changes following intranasal oxytocin in humans. *Biological Psychiatry*, 79(8), 693–705. doi: 10.1016/j.biopsych.2014.10.005

Parker, K. J., Oztan, O., Libove, R. A., Mohsin, N., Karhson, D. S., Sumiyoshi, R. D., … Hardan, A. Y. (2019). A randomized placebo-controlled pilot trial shows that intranasal vasopressin improves social deficits in children with autism. *Science Translational Medicine*, 11(491), eaau7356. doi: 10.1126/scitranslmed.aau7356

Pincus, D., Kose, S., Arana, A., Johnson, K., Morgan, P. S., Borckardt, J., … Nahas, Z. (2010). Inverse effects of oxytocin on attributing mental activity to others in depressed and healthy subjects: A double-blind placebo controlled fMRI study. *Frontiers in Psychiatry*, 1, 134. doi: 10.3389/fpsyt.2010.00134

Quintana, D. S., Lischke, A., Grace, S., Scheele, D., Ma, Y., & Becker, B. (2021). Advances in the field of intranasal oxytocin research: Lessons learned and future directions for clinical research. *Molecular Psychiatry*, 26, 80–91. doi: 10.1038/s41380-020-00864-7

Radke, S., & de Bruijn, E. R. A. (2015). Does oxytocin affect mind-reading? A replication study. *Psychoneuroendocrinology*, 60, 75–81. doi: 10.1016/j.psyneuen.2015.06.006

Riem, M. M. E., Bakermans-Kranenburg, M. J., Voorthuis, A., & van IJzendoorn, M. H. (2014). Oxytocin effects on mind-reading are moderated by experiences of maternal love withdrawal: An fMRI study. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 51, 105–112. doi: 10.1016/j.pnpbp.2014.01.014

Rilling, J. K., DeMarco, A. C., Hackett, P. D., Chen, X., Gautam, P., Stair, S., … Pagnoni, G. (2014). Sex differences in the neural and behavioral response to intranasal oxytocin and vasopressin during human social interaction. *Psychoneuroendocrinology*, 39, 237–248. doi: 10.1016/j.psyneuen.2013.09.022

Rilling, J. K., DeMarco, A. C., Hackett, P. D., Thompson, R., Ditzen, B., Patel, R., & Pagnoni, G. (2012). Effects of intranasal oxytocin and vasopressin on cooperative behavior and associated brain activity in men. *Psychoneuroendocrinology*, 37, 447–461. doi: 10.1016/j.psyneuen.2011.07.013

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin and Review*, 16(2), 225–237. doi: 10.3758/PBR.16.2.225

Sabatinelli, D., Fortune, E. E., Li, Q., Siddiqui, A., Krafft, C., Oliver, W. T., … Jeffries, J. (2011). Emotional perception: Meta-analyses of face and natural scene processing. *NeuroImage*, 54(3), 2524–2533. doi: 10.1016/j.neuroimage.2010.10.011

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience and Biobehavioral Reviews*, 42, 9–34. doi: 10.1016/j.neubiorev.2014.01.009

Schwaiger, M., Heinrichs, M., & Kumsta, R. (2019). Oxytocin administration and emotion recognition abilities in adults with a history of childhood adversity. *Psychoneuroendocrinology*, 99, 66–71. doi: 10.1016/j.psyneuen.2018.08.025

Spengler, F. B., Schultz, J., Scheele, D., Essel, M., Maier, W., Heinrichs, M., & Hurlemann, R. (2017). Kinetics and dose dependency of intranasal oxytocin effects on amygdala reactivity. *Biological Psychiatry*, 82(12), 885–894. doi: 10.1016/j.biopsych.2017.04.015

Spunt, R. P., & Adolphs, R. (2014). Validating the Why/How contrast for functional MRI studies of Theory of Mind. *NeuroImage*, 99, 301–311. doi: 10.1016/j.neuroimage.2014.05.023

Spunt, R. P., Falk, E. B., & Lieberman, M. D. (2010). Dissociable neural systems support retrieval of how and why action knowledge. *Psychological Science*, 21(11), 1593–1598. doi: 10.1177/0956797610386618

Spunt, R. P., & Lieberman, M. D. (2012a). An integrative model of the neural systems supporting the comprehension of observed emotional behavior. *NeuroImage*, 59(3), 3050–3059. doi: 10.1016/j.neuroimage.2011.10.005

Spunt, R. P., & Lieberman, M. D. (2012b). Dissociating modality-specific and supramodal neural systems for action understanding. *Journal of Neuroscience*, 32(10), 3575–3583. doi: 10.1523/JNEUROSCI.5715-11.2012

Spunt, R. P., Meyer, M. L., & Lieberman, M. D. (2015). The default mode of human brain function primes the intentional stance. *Journal of Cognitive Neuroscience*, 27(6), 1116–1124. doi: 10.1162/jocn_a_00785

Spunt, R. P., Satpute, A. B., & Lieberman, M. D. (2011). Identifying the what, why, and how of an observed action: An fMRI study of mentalizing and mechanizing during action observation. *Journal of Cognitive Neuroscience*, 23(1), 63–74. doi: 10.1162/jocn.2010.21446

Striepens, N., Kendrick, K. M., Hanking, V., Landgraf, R., Wüllner, U., Maier, W., & Hurlemann, R. (2013). Elevated cerebrospinal fluid and blood concentrations of oxytocin following its intranasal administration in humans. *Scientific Reports*, 3, 3440. doi: 10.1038/srep03440

Sun, R., Vuillier, L., Deakin, J., & Kogan, A. (2020). Oxytocin increases emotional theory of mind, but only for low socioeconomic status individuals. *Heliyon*, 6(3), e03540. doi: 10.1016/j.heliyon.2020.e03540

Tabak, B. A., Meyer, M. L., Castle, E., Dutcher, J. M., Irwin, M. R., Han, J. H., … Eisenberger, N. I. (2015). Vasopressin, but not oxytocin, increases empathic concern among individuals who received higher levels of paternal warmth: A randomized controlled trial. *Psychoneuroendocrinology*, 51, 253–261. doi: 10.1016/j.psyneuen.2014.10.006

Tabak, B. A., Meyer, M. L., Dutcher, J. M., Castle, E., Irwin, M. R., Lieberman, M. D., & Eisenberger, N. I. (2016). Oxytocin, but not vasopressin, impairs social cognitive ability among individuals with higher levels of social anxiety: A randomized controlled trial. *Social Cognitive and Affective Neuroscience*, 11(8), 1272–1279. doi: 10.1093/scan/nsw041

Tabak, B. A., Teed, A. R., Castle, E., Dutcher, J. M., Meyer, M. L., Bryan, R., … Eisenberger, N. I. (2019). Null results of oxytocin and vasopressin administration across a range of social cognitive and behavioral paradigms: Evidence from a randomized controlled trial. *Psychoneuroendocrinology*, 107, 124–132. doi: 10.1016/j.psyneuen.2019.04.019

Teed, A. R., Han, K., Rakic, J., Mark, D. B., & Krawczyk, D. C. (2019). The influence of oxytocin and vasopressin on men's judgments of social dominance and trustworthiness: An fMRI study of neutral faces. *Psychoneuroendocrinology*, 106, 252–258. doi: 10.1016/j.psyneuen.2019.04.014

Thijssen, S., Van 't Veer, A. E., Witteman, J., Meijer, W. M., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2018). Effects of vasopressin on neural processing of infant crying in expectant fathers. *Hormones and Behavior*, 103, 19–27. doi: 10.1016/j.yhbeh.2018.05.014

Tomova, L., Heinrichs, M., & Lamm, C. (2019). The other and me: Effects of oxytocin on self-other distinction. *International Journal of Psychophysiology*, 136, 49–53. doi: 10.1016/j.ijpsycho.2018.03.008

Uzefovsky, F., Shalev, I., Israel, S., Knafo, A., & Ebstein, R. P. (2012). Vasopressin selectively impairs emotion recognition in men. *Psychoneuroendocrinology*, 37(4), 576–580. doi: 10.1016/j.psyneuen.2011.07.018

Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *NeuroImage*, 48(3), 564–584. doi: 10.1016/j.neuroimage.2009.06.009

Walum, H., Waldman, I. D., & Young, L. J. (2016). Statistical and methodological considerations for the interpretation of intranasal oxytocin studies. *Biological Psychiatry*, 79(3), 251–257. doi: 10.1016/j.biopsych.2015.06.016

Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012). Conn: A functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connectivity*, 2(3), 125–141. doi: 10.1089/brain.2012.0073

Woolley, J. D., Chuang, B., Lam, O., Lai, W., O'Donovan, A., Rankin, K. P., … Vinograd, S. (2014). Oxytocin administration enhances controlled social cognition in patients with schizophrenia. *Psychoneuroendocrinology*, 47, 116–125. doi: 10.1016/j.psyneuen.2014.04.024

Zink, C. F., & Meyer-Lindenberg, A. (2012). Human neuroimaging of oxytocin and vasopressin in social cognition. *Hormones and Behavior*, 61(3), 400–409. doi: 10.1016/j.yhbeh.2012.01.016