SUPPLEMENTARY MATERIALS

for

Oxytocin attenuates trust as a subset of more general reinforcement learning, with altered reward circuit functional connectivity in males.

Jaime S. Ide, Sanja Nedic, Kin F. Wong, Shmuel L. Strey, Elizabeth A. Lawson, Bradford C. Dickerson, Lawrence L. Wald, Giancarlo La Camera, L.R. Mujica-Parodi

Detailed Materials and Methods

We conducted randomized, placebo-controlled crossover experiments using a single dose of intranasal OT (40 IU), in a cohort of subjects recruited through the project "*Computational modeling of oxytocin in the regulation of trust.*" The protocols described here were approved by the Institutional Review Boards of Stony Brook University and Partners HealthCare. The study was registered as Clinical Trial NCT01834261.

Subjects and Screening Procedures. Seventeen healthy male subjects (mean age 25.4 ± 3.7 years, weight 74 ± 10 kg) were tested at the MGH/HST Martinos Center for Biomedical Imaging. After an initial phone screening, a study physician obtained written informed consent from each subject. In addition, all subjects underwent History and Physical (H&P) exam with one of the study physicians or an MGH Clinical Research Center (CRC) nurse practitioner under the supervision of a study physician. Exclusion criteria included any known medical conditions (including mental disorders), metal in the body, claustrophobia, current use of any type of psychotropic medication, body mass index greater than 30, blood pressure above the normal range (140/90 mm Hg) or controlled with medication, smoking, and nasal obstruction. Subjects were instructed to abstain from caffeine and alcohol on the day of the scan.

Oxytocin/Placebo Procedures. Syntocinon (Oxytocin) Nasal Spray (manufactured by Novartis) was administered under IND # 112931. A registered nurse provided by the CRC, under supervision of a study physician, was present during drug administration and during each MRI

scanning session. Subjects received 5 sprays of oxytocin per nostril, alternating between nostrils, for a total of 10 sprays (40IU, 1mL), 60 minutes prior to fMRI scanning. Placebo, identical in preparation except for the oxytocin component, was administered in the same manner in a double blind, single-dose, randomized procedure. The MGH research pharmacy assigned drug randomization. To avoid bleed-through between conditions while controlling for order effects, each session was either Oxytocin (OT) only or Placebo (PL) only. All subjects underwent two scanning sessions (OT and PL) on separate days. Within subject, the two drugs (OT and PL) were administered at the same time of the day (within 2 hours), to control for possible diurnal variations in endogenous levels of OT.

Magnetic Resonance Imaging. Scanning was performed using a 7T Siemens Magnetom MRI scanner with a 32-channel headcoil. Each scanning session (OT and PL) included two fMRI scans: a 10-minute behavioral task scan (starting 60 minutes post drug administration, described below, and in **Fig. 2**), followed by a 10-minute resting state scan. Fixation cross was presented during resting state and subjects were instructed to keep their eyes open and refrain from any structured thoughts. In addition to BOLD fMRI data, we acquired Field Map images (used to correct distortions in EPI BOLD data) and T1-weighted structural (MEMPRAGE) images (coregistered with functional images and used in segmentation procedures). Whole brain EPI BOLD data were acquired using an optimized protocol with the following parameters: SMS slice acceleration factor = 5, R = 2 acceleration in the primary phase encoding direction (62 reference lines) and online GRAPPA image reconstruction, TR = 802 ms, TE = 20 ms, flip angle = 33°, voxel size = 2 x 2 x 1.5mm, slices = 85, number of measurements = 748, for a total acquisition time of 10 minutes 29 seconds. Field Map images were acquired using the following parameters: TR = 723 ms, TE 1 = 4.60 ms, TE 2 = 5.62 ms, flip angle = 36°, voxel size = 1.7 x 1.7 x 1.5 mm, slices = 89, for a total acquisition time of 3 minutes 14 seconds. The whole-brain T1-weighted structural volumes were acquired using a conventional multi-echo MPRAGE (MEMPRAGE) sequence with 1 mm isotropic voxel size and four echoes with the following protocol parameters: TE1 = 1.61ms, TE2 = 3.47ms, TE3 = 5.33ms, TE4 = 7.19ms, TR = 2530ms, flip angle = 7°, with R = 2 acceleration in the primary phase encoding direction (32 reference lines) and online GRAPPA image reconstruction, for a total acquisition time of 6 min 3 seconds per volume.

Behavioral Task (Iterative Trust Game).

To keep subjects engaged in the task, they were compensated for the participation in the study in the amount proportional to the amount of money earned during the game ($20-$50 per game). Prior to the first scanning session, subjects were trained (outside of the scanner) to play the game by completing two rounds of the game with the study coordinator acting as trustee. During the task, the subject was told that he/she was playing against another human, and the deception was only revealed upon the completion of the study. Unlike in King-Casas et al. (King-Casas et al., 2005), in which there were two human players, this deception was necessary in order to keep the social context of the game. The task was programmed in MATLAB. Each round consisted of the following periods: cue to invest, investment period, delay, investment reveal, delay, cue to repay, repayment period, delay, repayment reveal, delay, totals reveal, and inter-round delay. Timeline of the game and exact durations of periods are shown in **Fig. 1**. Delay periods had variable durations, optimized for this task using Optseq (https://surfer.nmr.mgh.harvard.edu/optseq/), varying between 2 and 7 seconds. The algorithm used for "trustee" repayment aimed to mimic real subjects, by simulating the malevolent and benevolent behaviors observed in the previous study by King-Casas et al. (King-Casas et al., 2005) (personal communication with Montague's group was made to access behavioral data from 48 subjects performing the repeated trust game). We used a fixed set of rules based on investor's decision, with parameters estimated from the real data, to simulate the trustee's repayments, with a degree of randomness of 10%.

Bayesian expectations of trust. As in our previous work (Ide et al., 2015; Ide et al., 2013), we used a dynamic Bayesian model (Yu and Cohen, 2009) to estimate the posterior belief of trust P(trust) during the repeated trust game. P(trust) values were computed on each event (trial) after a "trust signal", based on previous event history and current observation. We assumed that trust signal was "1" whenever the *investment ratio* (investment divided by 20) was increased (see (King-Casas et al., 2005) for details) or the repayment was larger than investment, and it was "0" otherwise. The Bayesian model assumed that the subject's belief in the trustee (trust) r_k on trial k has probability θ of being the same as r_k-1, and probability (1-θ) of being re-sampled from a fixed distribution $\pi(r_k)$. Subjects were also assumed to believe that trial k has probability r_k of signaling trust (s=1), and probability 1- r_k of not signaling trust (s=0). Based on these generative assumptions, subjects were assumed to use Bayesian inference to update their prior belief of

trusting on trial k, p($r_k|s_{k-1}$) based on the prior on the last trial p($r_{k-1}|s_{k-1}$) and last trial's true category ($s_k=1$ for trust signal, $s_k=0$ otherwise), where $\boldsymbol{s_k}=[s_1,…, s_k]$ is short-hand for all trials 1 through k. Specifically, given that the posterior distribution was p($r_{k-1}|s_{k-1}$) on trial k-1, the prior distribution of trust in trial k is given by: p($r_k|s_{k-1}$) = θ p($r_{k-1}|s_{k-1}$) + (1-θ) π(r_k), where the fixed distribution π(r_k) is assumed to be a beta distribution with prior mean pm and shape parameter "scale" sc, and the posterior distribution is computed from the prior distribution and the outcome according to the Bayes' rule: p($r_k|\boldsymbol{s_k}$) α p($s_k|r_k$) p($r_k|s_{k-1}$). The initial posterior distribution p($r_{k-1}|s_{k-1}$), for k=1, is assumed to be the same as the prior π(r_k). The likelihood function is p($s_k|r_k$) = r_k for s_k=1, and p($s_k|r_k$) = 1-r_k otherwise. Finally, the Bayesian estimate of trust on trial k, which we colloquially call P(trust), is simply the mean of the predictive distribution p($r_k|s_k$-1). Further details can be found in (Ide et al., 2015; Ide et al., 2013). In summary, for each subject, given a sequence of observed trust signals, and the three model parameters [θ, pm, sc], we estimated P(trust) for each trial. Generally speaking, parameter θ quantifies the weight given to the previous trials, and pm is the mean of the fixed belief on trustee. Importantly, these hidden measures of trust estimated trial-by-trial P(trust) were entered as parametric modulators in individual's general linear modeling, as done previously by us (Harle et al., 2014; Ide et al., 2013) and others (Daw et al., 2006; O'Doherty et al., 2006).

Reinforcement learning of trust. We used a standard reinforcement modeling (Dayan and Abbott, 2005; Sutton and Barto, 1998), similar to the implementation presented by Crone and colleagues (van den Bos et al., 2012). Given a set of possible actions $\{a_1,…,a_n\}$, and a set of associated action values $\{V_1,…,V_n\}$ where n is the number of possible choices, for each trial t, we update the action value of the currently selected choice j using the expression: V_j (t) = V_j (t-1) + α(r(t) – V_j (t-1)). α is the *learning rate*, and the difference r(t) – V_j (t-1) is the *prediction error* between the obtained reward r(t) and the expected action value V_j(t-1). Greater learning rate α designates greatest response to the reward feedback or prediction error. In this modeling, we assumed a single state and no decay parameter, i.e. γ=0 (subjects only consider the actual reward). Given the set of action values, we used a *softmax* decision policy (Dayan and Abbott, 2005; Sutton and Barto, 1998) without bias, where the probability associated with each choice a_j is computed using a sigmoid function $P(a_j) = \exp(\beta V_j) / \sum_{i=1}^{n} \exp(\beta V_i)$, where β is the inverse temperature.

Greater β indicates more deterministic greedy actions. For the repeated trust game, we defined the reward $r(t)$ as the total reward in each round t, and assumed that participants learn how much to trust as they update the value of each choice. In fact, in the repeated trust game, participants learn how much they can trust as they have an accurate prediction of their actions (low prediction errors). In practice, for each subject, given a sequence of observed actions and rewards, we found the optimal learning rates and inverse temperature that minimize the total negative log-likelihood, computed as the sum of $-\log\left(P\left(a_j|model\right)\right)$ of each observed action a_j. Since we considered models with the same number of parameters, we optimized the cost function in terms of the likelihood. We performed the optimization using the function *fminsearch* available in Matlab.

Preprocessing, GLM, and statistical analyses. The neuroimaging data were preprocessed and analyzed with Statistical Parametric Mapping 12 (SPM12) (Wellcome Department of Imaging Neuroscience, University College London, U.K.), as in our previous works (Ide et al., 2013; Ide et al., 2014; Li et al., 2014). Importantly, this pipeline was validated and demonstrated to be robust for processing ultra-high field MR images obtained at 7T scan. The anatomical images (MEMPRAGE) were normalized to an MNI (Montreal Neurological Institute) template using a unified segmentation and registration framework (Ashburner and Friston, 2005). Images of each individual subject were realigned to account for head movements, and unwarped to correct for geometric distortions caused by the enhanced magnet field inhomogeneity (Andersson et al., 2001; Hutton et al., 2002), followed by normalization to MNI space and smoothing (kernel of 6mm). One subject was excluded from neuroimaging analyses because of severe head motion, therefore all imaging group comparisons were performed with sixteen subjects (N=16). In the first level, standard general linear models (GLM) were constructed (Friston et al., 1995), using the experimental conditions as main regressors, six head-movement parameters as covariates, and entering P(trust) as parametric modulators (Daw et al., 2006; O'Doherty et al., 2006). All other statistical analyses were computed using MATLAB and Statistics Toolbox Release 2012b (The MathWorks, Inc., Natick, Massachusetts, United States).

Functional and effective connectivity analyses. To delineate functional interactions within the reward and learning brain networks, in a data-driven manner, we used psycho-physiological

interactions or PPI as previously applied by our group (Duann et al., 2009; Ide and Li, 2011). To define the seeds, we employed anatomically defined masks for amygdala, nucleus accumbens and OFC (Desikan et al., 2006; Tziortzi et al., 2014; Zaborszky et al., 2008) available in FSL (FMRIB Software Library v5.0). Dynamic causal modeling or DCM12 (Friston et al., 2003) was used to test competing brain-circuit hypothesis, defined by our electrophysiological studies in rats. Standard voxel of interest (VOI) time series extraction were performed by computing the first engenvariate inside the ROI masks and adjusting for effects of interest (Stephan et al., 2010). Bayesian model selection (BMS) were used to select the most likely model from a set of plausible models (Stephan et al., 2009; Stephan et al., 2010). We used DCM to delineate the circuit comprised of nucleus accumbens, amygdala, and OFC, regions traditionally affected by OT (Bethlehem et al., 2013). Once the best DCMs were selected and estimated, the individual parameters of intrinsic and modulatory connections were correlated with behavioral variables (Ballard et al., 2011). Additionally, we used the spectral DCM (spDCM), as opposed to traditional DCM, since it allows modeling endogenous or random fluctuations of intrinsic dynamics (Friston et al., 2014; Razi et al., 2015). Recently, it was shown that spDCM is not only comparable, but more accurate than stochastic DCM in modeling neuronal fluctuations of resting state fMRI data (Razi et al., 2015).
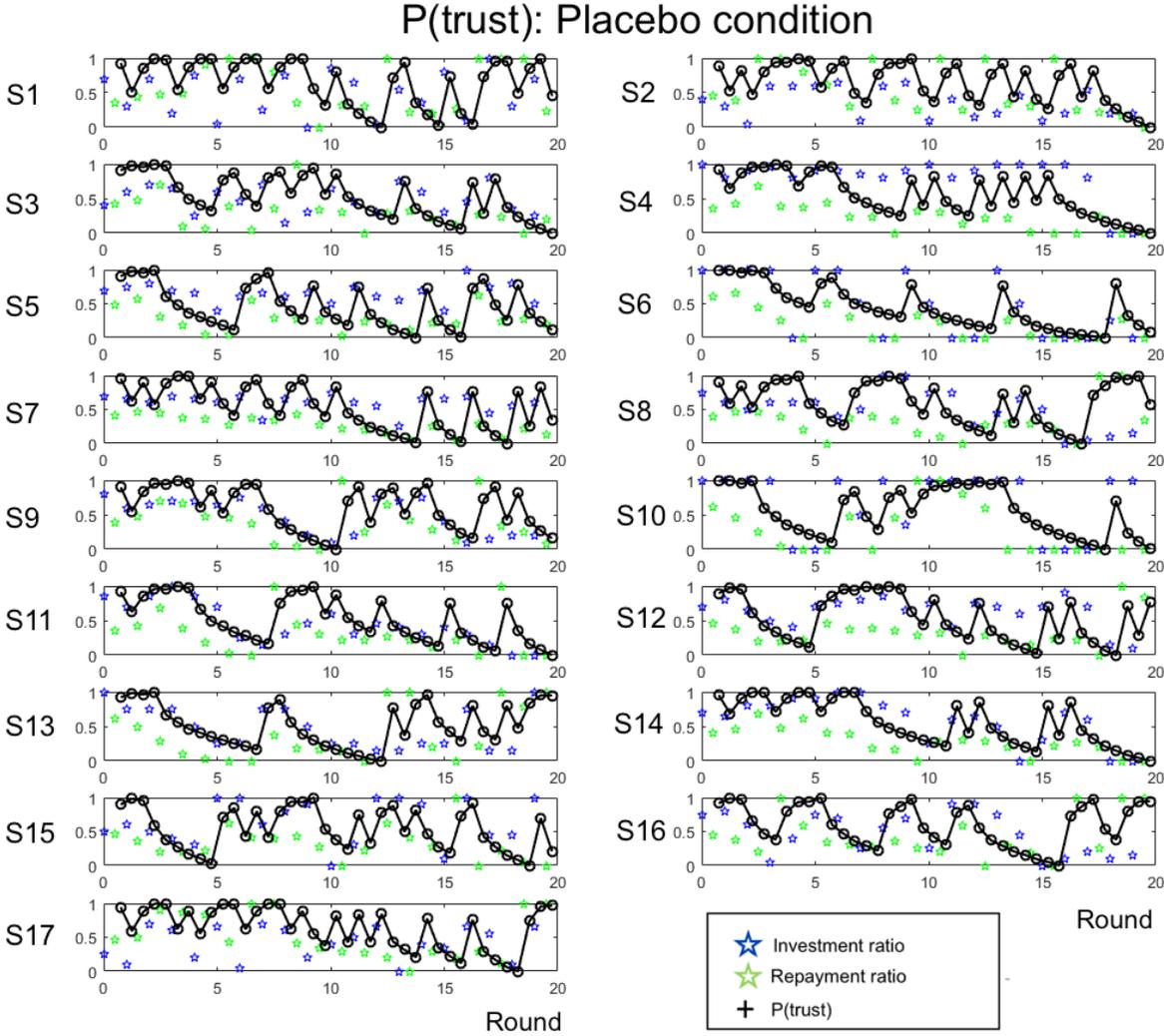
Additional Figures and Tables

Figure S1. **Estimated individual expected values of trust, P(trust), are correlated with the investment ratio.** Expected values of trust for each subject for the placebo condition.
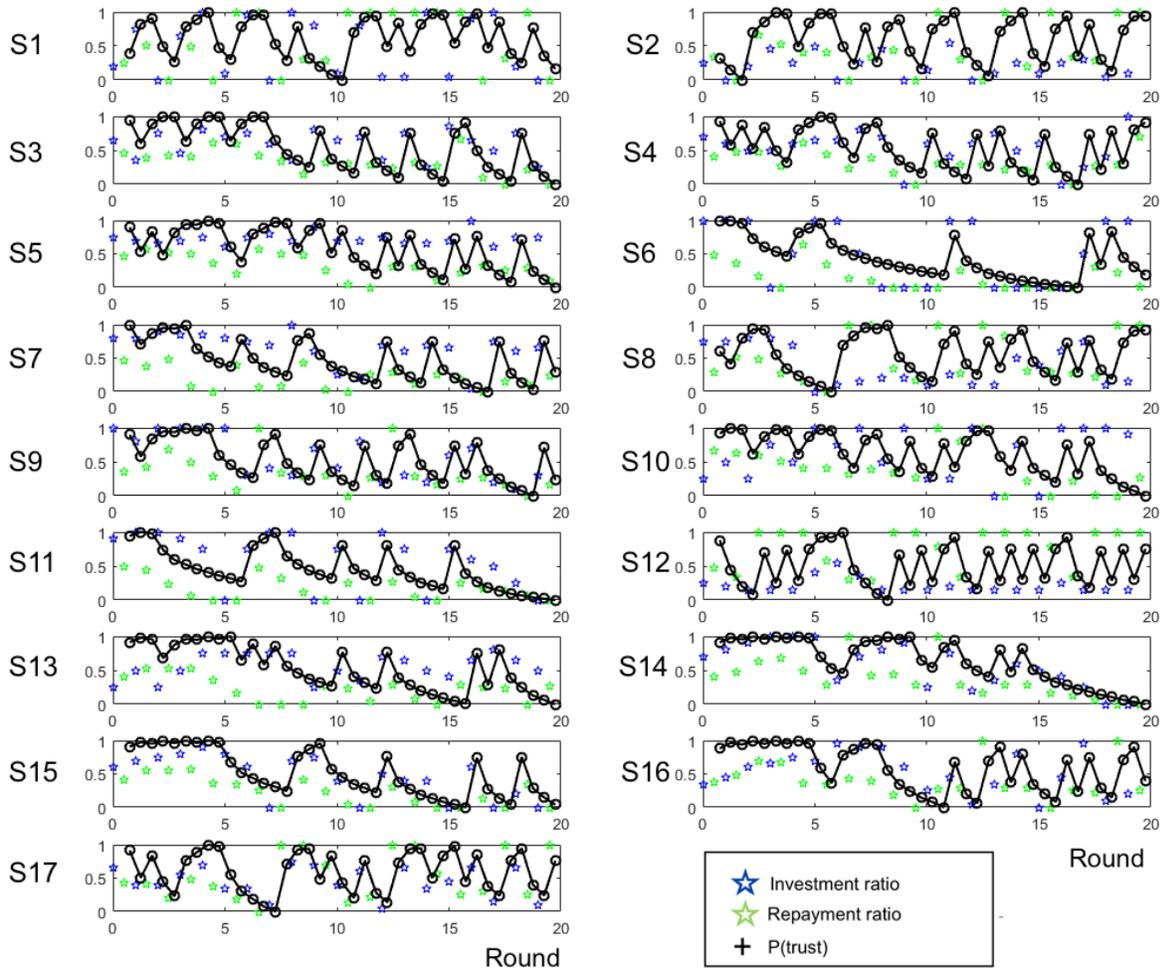
Figure S2. **Estimated individual expected values of trust, P(trust), show weaker correlations with the investment ratio under oxytocin.** Expected values of trust for each subject for the oxytocin condition.
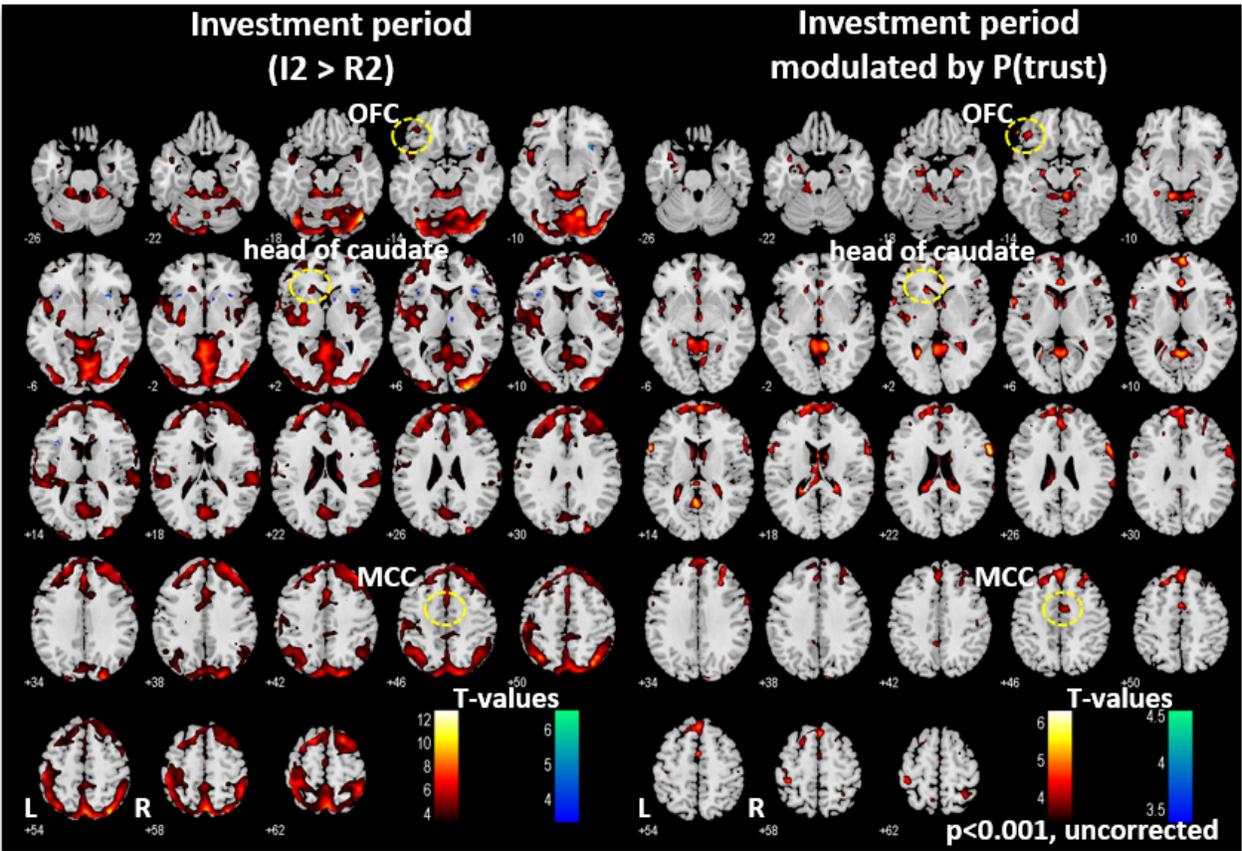
Figure S3. **Clusters in the OFC, head of caudate and MCC are activated during investment period and significantly modulated by P(trust)**. Brain regions activated during investment as compared to repayment period (I2 > R2), and modulated by P(trust).
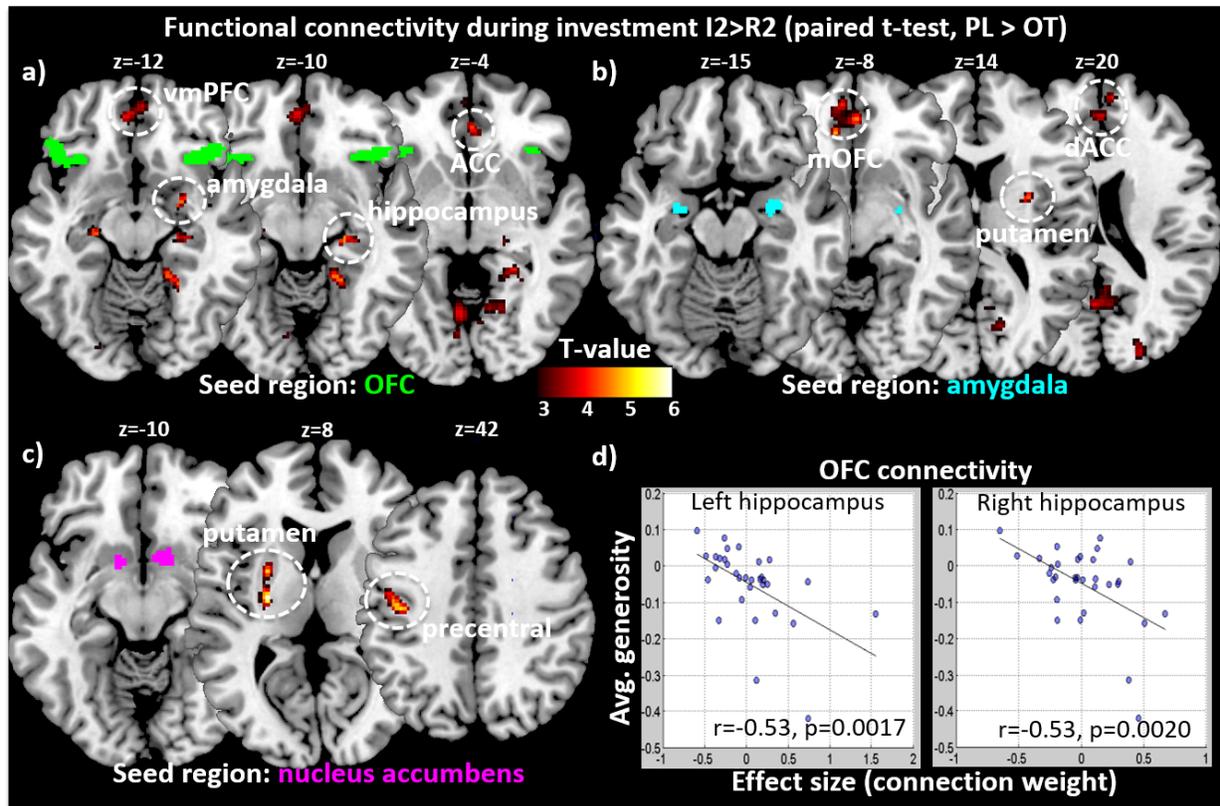
Figure S4. Brain regions with reduced functional connectivity during investment phase (I2>R2) for OT as compared to PL conditions (paired t-test, p<0.005, uncorrected). *(a)* Regions with reduced functional connectivity (seed region: OFC, green color) during investment period, for OT as compared to PL conditions. *(b)* Regions with reduced functional connectivity (seed region: bilateral amygdala, cyan color) during investment period, for OT as compared to PL conditions. *(c)* Regions with reduced functional connectivity (seed region: NAcc, magenta color) during investment period, OT>PL. *(d)* OFC|Hipp connectivity correlates with generosity.
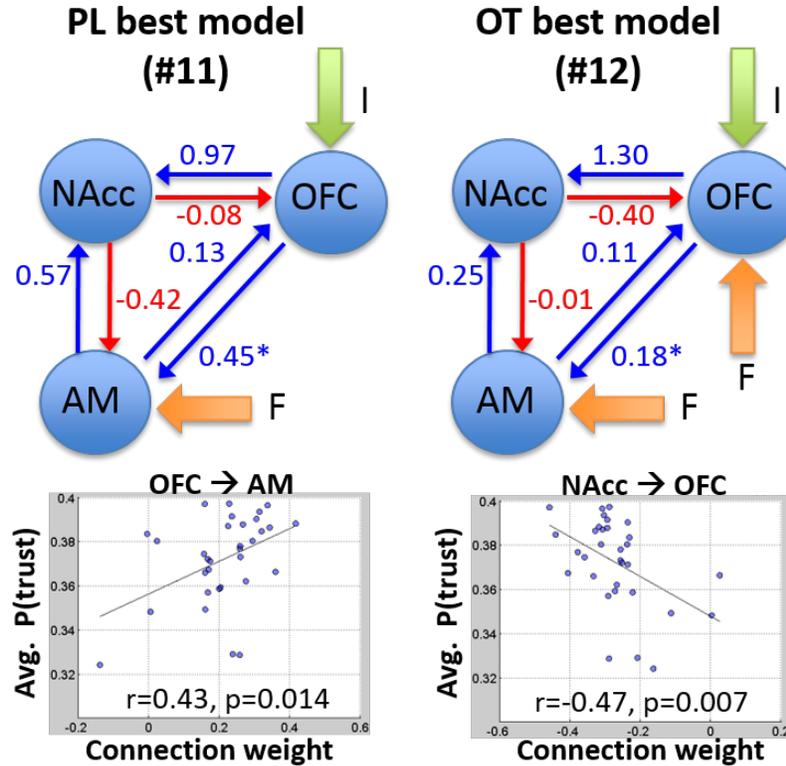
PL best model (#11)

0.97
NAcc → OFC
-0.08
0.13
0.57
-0.42
0.45*
AM
I
F

OFC → AM
r=0.43, p=0.014
Avg. P(trust)
Connection weight

OT best model (#12)

1.30
NAcc → OFC
-0.40
0.11
0.25
-0.01
0.18*
AM
I
F

NAcc → OFC
r=-0.47, p=0.007
Avg. P(trust)
Connection weight

Figure S5. OT reduces effective connectivity between OFC and amygdala. (Upper panel) Group winning models for PL and OT conditions are represented (Models #11 and #12, respectively), as well as the average group parameters after Bayesian parameter averaging (BPA) given the model. The average weight of OFC→AM is 0.18 for OT group, while it is 0.45 for PL group for their respective best models. **(Lower panel)** Connection weight of OFC→AM (Model #11) is positively correlated with average P(trust) (r=0.42, p=0.014), while connection weight of NAcc→OFC (Model #12) is negatively correlated with average P(trust) (r=-0.47, p=0.007). Driving inputs: investment "I" and feedback "F". Regions, NAcc: nucleus accumbens, AM: amygdala, OFC: orbitofrontal cortex.

Table S1. OT does not increase risk-taking. Inverse temperatures (β) were equivalent under OT ($\beta=0.30\pm0.29$) as compared to PL ($\beta=0.16\pm0.23$) conditions (paired t-test with bootstrapping, p=0.1720; after removing subjects with learning rates > 1 (#8 and #14)).

S#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
PL	0.03	0.06	0.10	0.19	0.15	0.11	0.16	n/a	0.12	0.14	0.05	0.14	0.01	n/a	0.95	0.07	0.07
OT	0.65	0.60	0.13	0.10	0.20	0.14	0.87	n/a	0.1	0.43	0.09	0.20	0.77	n/a	0.08	0.05	0.06

Supplementary References

Andersson JL, Hutton C, Ashburner J, Turner R, Friston K (2001) Modeling geometric deformations in EPI time series. Neuroimage 13:903-919.

Ashburner J, Friston KJ (2005) Unified segmentation. NeuroImage 26:839-851.

Ballard IC, Murty VP, Carter RM, MacInnes JJ, Huettel SA, Adcock RA (2011) Dorsolateral prefrontal cortex drives mesolimbic dopaminergic regions to initiate motivated behavior. The Journal of neuroscience : the official journal of the Society for Neuroscience 31:10340-10346.

Bethlehem RA, van Honk J, Auyeung B, Baron-Cohen S (2013) Oxytocin, brain physiology, and functional connectivity: a review of intranasal oxytocin fMRI studies. Psychoneuroendocrinology 38:962-974.

Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. Nature 441:876.

Dayan P, Abbott LF (2005) Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems. MIT Press.

Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ (2006) An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31:968-980.

Duann JR, Ide JS, Luo X, Li CS (2009) Functional connectivity delineates distinct roles of the inferior frontal cortex and presupplementary motor area in stop signal inhibition. The Journal of neuroscience : the official journal of the Society for Neuroscience 29:10171-10179.

Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. NeuroImage 19:1273.

Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ (1995) Statistical parametric maps in functional imaging: A general linear approach. Human Brain Mapping 2:189-210.

Friston KJ, Kahan J, Biswal B, Razi A (2014) A DCM for resting state fMRI. Neuroimage 94:396-407.

Harle KM, Shenoy P, Stewart JL, Tapert SF, Yu AJ, Paulus MP (2014) Altered neural processing of the need to stop in young adults at risk for stimulant dependence. The Journal of neuroscience : the official journal of the Society for Neuroscience 34:4567-4580.

Hutton C, Bork A, Josephs O, Deichmann R, Ashburner J, Turner R (2002) Image distortion correction in fMRI: A quantitative evaluation. Neuroimage 16:217-240.

Ide JS, Hu S, Zhang S, Yu AJ, Li CS (2015) Impaired Bayesian learning for cognitive control in cocaine dependence. Drug Alcohol Depend.

Ide JS, Li CSR (2011) Error-Related Functional Connectivity of the Habenula in Humans. Frontiers in Human Neuroscience 5.

Ide JS, Shenoy P, Yu AJ, Li CS (2013) Bayesian prediction and evaluation in the anterior cingulate cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience 33:2039-2047.

Ide JS, Zhang S, Hu S, Sinha R, Mazure CM, Li CS (2014) Cerebral gray matter volumes and low-frequency fluctuation of BOLD signals in cocaine dependence: Duration of use and gender difference. Drug Alcohol Depend 134:51-62.

King-Casas B, Tomlin D, Anen C, Camerer CF, Quartz SR, Montague PR (2005) Getting to know you: reputation and trust in a two-person economic exchange. Science 308:78-83.

Li CS, Ide JS, Zhang S, Hu S, Chao HH, Zaborszky L (2014) Resting state functional connectivity of the basal nucleus of Meynert in humans: in comparison to the ventral striatum and the effects of age. Neuroimage 97:321-332.

O'Doherty JP, Buchanan TW, Seymour B, Dolan RJ (2006) Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. Neuron 49:157-166.

Razi A, Kahan J, Rees G, Friston KJ (2015) Construct validation of a DCM for resting state fMRI. Neuroimage 106:1-14.

Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ (2009) Bayesian model selection for group studies. NeuroImage 46:1004-1017.

Stephan KE, Penny WD, Moran RJ, den Ouden HEM, Daunizeau J, Friston KJ (2010) Ten simple rules for dynamic causal modeling. NeuroImage 49:3099-3109.

Sutton RS, Barto AG (1998) Introduction to Reinforcement Learning. MIT Press.

Tziortzi AC, Haber SN, Searle GE, Tsoumpas C, Long CJ, Shotbolt P, Douaud G, Jbabdi S, Behrens TE, Rabiner EA, Jenkinson M, Gunn RN (2014) Connectivity-based functional analysis of dopamine release in the striatum using diffusion-weighted MRI and positron emission tomography. Cerebral cortex 24:1165-1177.

van den Bos W, Cohen MX, Kahnt T, Crone EA (2012) Striatum-medial prefrontal cortex connectivity predicts developmental changes in reinforcement learning. Cerebral cortex 22:1247-1255.

Yu A, Cohen J (2009) Sequential effects: Superstition or rational behavior? In: NIPS 2008. Koller D, Schuurmans D, Bengio Y, Bottou L (eds). MIT Press: Vancouver, British Columbia, Canada. pp 1873-1880.

Zaborszky L, Hoemke L, Mohlberg H, Schleicher A, Amunts K, Zilles K (2008) Stereotaxic probabilistic maps of the magnocellular cell groups in human basal forebrain. Neuroimage 42:1127-1141.