nature research

Corresponding author(s): Yina MA

Last updated by author(s): Jan 7, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Presentation 17.0, MATLAB 2014b
Data analysis	MATLAB 2016b,R Studio 1.1.414, SPM12

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are available upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The sample size of the current study was determined prior to data collection and was detailed as follows: 1. Sample size estimation for the fMRI experiment. The sample size of the fMRI study was determined prior to data collection. We conducted sample size estimation using G*Power 3.1 (Faul et al.,2007) to determine the number of participants sufficient to detect a reliable effect. Based on an estimated average small-to-medium effect size of oxytocin effect on social behaviors (Cohen's d = 0.28, Walum et al., 2016). 104 participants were needed to detect a significant effect ($\alpha = 0.05$, $\beta = 0.80$, two-by-two mixed ANOVA interaction effects). We planned to recruit 125 participants (assuming 10-20% participants would be removed from the fMRI data analysis due to excessive head movement). In the end, we recruited 127 participants because the 41th and 42th participants did not complete the experiment due to technical issues during scanning. For comparison, we also considered the 58 oxytocin-fMRI studies published at the time we initiated our experiment in June 2015, of which 23 employed between-subject design recruiting healthy individuals. On average, the sample size was 50.89 in total, 25.82 for the placebo group and 25.47 for the oxytocin group. Thus, our planned sample size of 125 participants was a decent sample size compared to the average across oxytocin-fMRI studies. Moreover, the sample size of 116 participants (after removal of subjects due to technical issues and excessive head movement) was adequate to reveal reliable effects, exceeding the 104 participants needed for 80% power. 2. Sample size estimation for the oxytocin replication experiment. We conducted an additional behavioral experiment for the replication of the oxytocin effect using a double-blind, randomized, placebo-controlled, within-subjects crossover design. The sample size was predetermined based on the effect size (Cohen's d = 0.45) from our original finding in the fMRI study. The G*Power calculation suggested that 40 participants (20 for each group) were required to detect a reliable effect ($\alpha = 0.05$, $\beta = 0.80$ for a within (oxytocin vs. placebo)-between (prosocial vs. individualist) interaction). To obtain a better sense of the robustness of the original findings, we doubled the estimated sample size, aiming to enroll 40 participants per group, with corresponding power equal to 98%. We replicated the selective oxytocin effects on promoting prosociality (i.e., ϕ) in individualists in the whole sample (40 prosocials and 40 individualists), as well as in the first 20 prosocials and 20 individualists (as estimated by the G*Power analysis). 140 males (mean age, 22.33 ± 3.35 years) were invited to a behavioral session to identify their disposition in social value orientation. Among these, 82 participants were qualified and willing to participate in the oxytocin experiment (at least 7 days after the first behavioral session). Two participants did not show up for the second session. Thus, 80 participants (40 prosocials, mean age, 22.08 ± 3.47 years; 40 individualists; mean age, 21.54 ± 2.42 years) were included in the final data analysis. Reference. Faul, F., Erdfelder, E., Lang, A. G. & Buchner, A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res methods 39, 175-191 (2007). Walum, H., Waldman, I.D. & Young, L.J. Statistical and Methodological Considerations for the Interpretation of Intranasal Oxytocin Studies. Biol. Psychiatry. 79, 251-257 (2016).
Data exclusions	The exclusion criteria were established prior to data analysis based on previously published papers in the field. Subjects with poor performance and excessive head motion during scanning would be excluded, which was in line with with a majority of fMRI studies. 1. Behavior: Button responses were recorded for each trial. Responses that were not recorded within four seconds were discarded, based on the time limit of jitter. 2. fMRI: Scans that exceeded 3 mm of intravolume motion were excluded in line with Power et al.,2012's instruction. Reference. Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage, 59, 2142-2154 (2012).
Replication	1. We had a post-scan behavioral experiment on the same participants went through fMRI studies - the results are largely the same. 2. Before the actual experiment, we have recruited another independent sample of participants (n=60) to perform a very similar task, the results are very similar as the main finding. 3. We have run two additional experiments (N = 315 and N = 140) with a revised experimental design. The first additional study with a large sample size (N = 315) provides a replication for our winning social referent point model. The second additional study was an oxytocin experiment that replicated the selective oxytocin effect in prosocials and individualists and the winning model again (N = 80 males for a within-subject, placebo-controlled design with oxytocin challenge, 40 prosocials and 40 individualists, 140 males screened for social value orientation). The two additional experiments were both run on a modified design that improved the ability to distinguish between different models. We replicated the effects reported in the original manuscript across both samples using the improved design. All results are reported in the main text.
Randomization	Participants recruited in the fMRI experiment were randomly assigned to the intranasal administration of oxytocin or placebo in a double-blind placebo-controlled between-subjects design. Oxytocin-replication experiment. We conducted an additional behavioral experiment for the replication of the oxytocin effect using a double-blind, randomized, placebo-controlled, within-subjects crossover design.The order of oxytocin and placebo treatment was counterbalanced across participants.
Blinding	The oxytocin-fMRI and oxytocin-replication experiments were double-blind, i.e., both participants and experimenters were blind to experimental conditions (both treatment and social disposition conditions). Data analysis was not performed blind to the conditions of the experiments.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	Oxytocin fMRI experiment: There were 282 male college students (mean age=22.3 ± 2.12 years) that participated in this study as paid volunteers. 127 participants were qualified and willing to participate in the fMRI experiment. Two participants (1.6%) were excluded due to technical issues during scanning, leaving 125 participants in the behavioral analysis. An additional 9 participants (7.2%) were excluded from further fMRI analysis due to excessive head movement during scanning (> 3 mm), leaving 116 participants for fMRI data analysis. In the end, there were 60 prosocials including 30 administered placebo (mean age, 22.7 ± 2.61 years) and 30 administered oxytocin (mean age, 21.8 ± 2.38 years), and 56 individualists including 30 administered placebo (mean age, 23.0 ± 2.29 years) and 26 administered oxytocin (mean age, 22.5 ± 3.03 years) in the formal fMRI data analysis. Prosocials and individualists receiving oxytocin or placebo were matched on state and trait anxiety, depression, subjective well-being and happiness ratings (all $p > 0.05$ on both the main and interaction effects of Treatment and Social Disposition; Supplementary Table 3a). Oxytocin-replication experiment: 140 males (mean age, 22.33 ± 3.35 years) were invited to a behavioral session to identify their disposition in social value orientation. Among these, 82 participants were qualified and willing to participate in the oxytocin experiment (at least 7 days after the first behavioral session). Two participants did not show up for the second session. Thus, 80 participants (40 prosocials, mean age, 22.08 ± 3.47 years; 40 individualists; mean age, 21.54 ± 2.42 years) were included in the final data analysis. Prosocials and individualists were matched on state and trait anxiety, depression, subjective well-being and happiness ratings (all $p > 0.05$ on both the main and interaction effects of Treatment and Social Disposition; Supplementary Table 3b). Online-replication experiment: We conducted an online experiment with a large sample (n = 315, 132 males, 160 prosocials, mean age = 22.40 ± 3.27; 155 individualists, mean age = 22.48 ± 3.30) to provide a replication for our finding that the social reference model outperforms other models. (82 female, age=22.40±3.27, SVO score=9.16±7.74) and155 individualists(101 female, age=22.48±3.30, SVO score= 33.62±5.63). Prosocials and individualists did not differ in their ratings on the first impression, likeability and attractiveness of the online partner (independent-samples t test, impression: $t313 = 1.03$, $p = 0.305$; likeability: $t313 = 0.98$, $p = 0.328$; attractiveness: $t313 = 0.73$, $p = 0.465$).
Recruitment	Oxytocin fMRI experiment: 282 male college students were recruited in this study as paid volunteers through on campus flyer recruitment. We first measured participant's disposition in social value orientation in the behavioral session. Among these, 127 participants were qualified and willing to participate in the fMRI experiment. Oxytocin-replication experiment: We recruited 140 males through on campus flyer recruitment to a behavioral session to identify their disposition in social value orientation. Among these, 82 participants were qualified and willing to participate in the oxytocin experiment. Online-replication experiment: This is an on-line behavioral study with 315 participants, which were recuited on Qualtrics. Although previous work has found demographic effects of recruiting volunteers into experiments that they may have less income and more leisure time (Cleave et al.,2013), our subject recruitment was based on the scores of social value orientation of randomly recruited particiants. Moreover, the within-subject design of the oxytocin replication experiment, as well as the randomization process also minimize the influence of population's difference. Ref. Cleave, B. L., Nikiforakis, N., & Slonim, R. Is there selection bias in laboratory experiments? The case of social and risk preferences. Experimental Economics, 16, 372-382 (2013).
Ethics oversight	The experimental protocol was in line with the standards of the Declaration of Helsinki and approved by the research ethics committee at the State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University (Beijing, China).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Magnetic resonance imaging

Experimental design

Design type	task fMRI, event-related design
Design specifications	2 scanning sessions per participant; 90 trials per session; trial length ranged from 7 to 12 seconds; inter-trial interval length ranged from 4 to 6 seconds.

Behavioral performance measures	Participants evaluated their preferences for a monetary allocation by button press in each trial. Button responses were recorded for each trial. Responses that were not recorded within four seconds were discarded. The preferences and decision times were measured.

Acquisition

Imaging type(s)	functional, structural
Field strength	3 Tesla
Sequence & imaging parameters	Whole-brain imaging data was collected on a GE 3-Tesla MR scanner with a standard head coil (HDx, Signa MR 750 System; GE Healthcare, Milwaukee, WI). Functional images were collected using an echo-planar imaging sequence (axial slices, 32; slice thickness, 4 mm; gap, 1 mm; TR, 2000 ms; TE, 30 ms; voxel size, 3.75 × 3.75 × 5 mm; flip angle, 90°; FOV, 240 × 240 mm; and 285 volumes for each session, two sessions in total). Structural images were acquired through 3D sagittal T1-weighted magnetization-prepared rapid gradient echo (180 slices; TR, 8.208 ms; TE, 3.22 ms; slice thickness, 1 mm; voxel size, 0.47 × 0.47 × 1.0 mm3; flip angle, 12°; inversion time, 450 ms; FOV, 240 × 240 mm).
Area of acquisition	Whole brain; each volume comprised 32 axial slices collected in an interleaved-ascending manner and parallel to the AC-PC line.

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software	Brain imaging data was preprocessed using Statistical Parametric Mapping (SPM12; http://www.fil.ion.ucl.ac.uk/spm). The first 5 functional images from each session were discarded for signal equilibrium and participants' adaptation to scanning noise. Remaining images were corrected for slice acquisition timing and realigned for head motion correction. Subsequently, functional images were coregistered to each participant's grey matter image segmented from corresponding high-resolution T1-weighted image, then spatially normalized into a common stereotactic Montreal Neurological Institute (MNI) space and resampled into 2-mm isotropic voxels. Finally, images were smoothed by an isotropic 3D Gaussian kernel with 8-mm full-width at half-maximum.
Normalization	We used the 'Old normalise' template from SPM12 which, according to the manual (http://www.fil.ion.ucl.ac.uk/spm/doc/manual.pdf), states the following: "The first step of the normalisation is to determine the optimum 12-parameter affine transformation. Initially, the registration is performed by matching the whole of the head (including thescalp) to the template. Following this, the registration proceeded by only matching the brains together, by appropriate weighting of the template voxels. This is a completely automated procedure (that does not require "scalp editing') that discounts the confounding effects of skull and scalp differences. A Bayesian framework is used, such that the registration searches for the solution that maximises the a posteriori probability of it being correct. i.e., it maximises the product of the likelihood function (derived from the residual squared difference) and the prior function (which is based on the probability of obtaining a particular set of zooms and shears). The affine registration is followed by estimating nonlinear deformations, whereby the deformations are defined by a linear combination of three dimensional discrete cosine transform (DCT) basis functions. The default options result in each of the deformation fields being described by 1176 parameters, where these represent the coefficients of the deformations in three orthogonal directions. The matching involved simultaneously minimising the membrane energies of the deformation fields and the residual squared difference between the images and template(s)."
Normalization template	Images were transformed to conform to the default T1 Montreal Neurological Institute (MNI) brain interpolated to 3 × 3 × 3 mm.
Noise and artifact removal	Scans that exceeded 3 mm of intravolume motion were excluded
Volume censoring	We used the ArtRepair toolbox to remove artifacts. Values for repaired scans were imputed and deweighted by interpolating between the nearest non-repaired scans.

Statistical modeling & inference

Model type and settings	mass univariate; level 1: fixed effects, high-pass filter with cutoff period of 128 s was used; level 2: random effects
Effect(s) tested	At first level, we created parametric regressors that were associated with the value distance or value difference between self and other, at the monetary outcome-pair presentation to search for brain regions that encoded the subjective distance in social value representation. In the fMRI analysis, we included 10 GLM models: a social value distance from an individual-specific reference-point (i.e., 1 - cos (θ (t) - ϕ), θ (t) is the angle of a potential allocation at trial t and ϕ is the individual-specific reference-point derived from our social reference model), an egocentric reference (i.e., cos (θ (t)), an allocentric reference (i.e., sin (θ(t)), an objective equality reference-point (i.e., cos (θ (t) - 45°), monetary outcome for the self (i.e., $Self), monetary outcome for the partner (i.e., $Other), absolute value difference (| $Self - $Other| or advantageous (i.e., max (0, $Self -$Other)) or disadvantageous inequality aversion (i.e., max (0, $Other -$Self)) separately, or with preference rating as parametric modulator in the GLM. At second level, Coefficients for each regressor were estimated for each participant using maximum likelihood estimates to account for serial correlations in the data. Statistical significance was determined at the group level using a random-effects analysis. Significant clusters from all 2nd-level analyses were determined using a height threshold of $P < 0.001$ and an extent threshold of $P < 0.05$ with family-wise error (FWE) correction. We also applied voxelwise inference

using FWE-corrected threshold of P = 0.05 on the whole-brain analysis, given recently recognized concern over clusterwise inference. For the relationship between value distance (1-cos(θ(t) – ϕ) and neural responses during monetary outcome pair presentation, the peak voxels in right amygdala survived voxelwise FWE correction (P = 0.02).

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☒ Both

Anatomical location(s)

All ROI analyses were employed on regions defined anatomically. Amygdala ROIs are defined based on AAL bilateral anatomical mask. LOFC are defined based on based on Neubert et al. (2014), combining connectivity-based parcellations 8-11, which includes all of right lOFC.
Ref. Neubert, F.X., Mars, R.B., Thomas, A.G., Sallet, J. & Rushworth, M.F. Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex. Neuron 81, 700-713 (2014).

Statistic type for inference
(See Eklund et al. 2016)

Statistical significance was determined at the group level using a random-effects analysis. Significant clusters from second-level analyses were determined using a height threshold of P < 0.001 and an extent threshold of P < 0.05 with cluster-based family-wise error (FWE) correction. We also applied voxelwise inference using the FWE-corrected threshold of P = 0.05 on the whole-brain analysis, given recent concern over cluster-wise inferences. For the relationship between value distance (1-cos(θ(t) – ϕ) and neural responses during monetary outcome-pair presentation, the peak voxels in right amygdala survived voxelwise FWE correction (P = 0.02).

Correction

Significant clusters from second-level analyses were determined using a height threshold of P < 0.001 and an extent threshold of P < 0.05 with cluster-based family-wise error (FWE) correction. We also applied voxelwise inference using the FWE-corrected threshold of P = 0.05 on the whole-brain analysis, given recent concern over cluster-wise inferences. For the relationship between value distance (1-cos(θ(t) – ϕ) and neural responses during monetary outcome-pair presentation, the peak voxels in right amygdala survived voxelwise FWE correction (P = 0.02).

Models & analysis

n/a	Involved in the study
☐	☒ Functional and/or effective connectivity
☒	☐ Graph analysis
☒	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Functional connectivity was measured. We performed a generalized PPI analysis with anatomically defined bilateral amygdala as the seed region at the whole-brain level.