Supplementary Information

Title: Oxytocin induces long-lasting adaptations within amygdala circuitry in autism: a treatment-mechanism study with randomized placebo-controlled design

Authors: Kaat **Alaerts**, Sylvie **Bernaerts**, Jellina **Prinsen**, Claudia **Dillen**, Jean **Steyaert**, Nicole **Wenderoth**

Supplementary Methods

Intervention

Administration instructions. All participants administered the first dose in front of the experimenter and received clear instructions about the use of the nasal spray in accordance with recommendations by Guastella et al. (2013) (Guastella et al. 2013). At first use, air present in the nasal spray was removed by pumping the spray until a fine mist was observed. Participants were instructed to keep one nostril closed, to take a deep breath through the nose and to tilt their head slightly backwards during nasal administration in order to minimize gravitational loss of the spray. All participants were monitored onsite until approximately two hours after first nasal spray administration.

Participants were asked to take the nasal spray in the morning; and to keep a daily record of the time point of nasal spray administration and whether or not they were alone or in company of others the first two hours after administration. Percentage of days at which the spray was administered in the presence of others was not significantly different between treatment groups (OT: 36.6 % (SD 29.9); PL: 37.4 % (SD 24.1); $t(36)= -.09$; $p= .92$). Used sprays were returned by 79.0% of the participants to analyze the amount of spray used (17 OT/ 13 PL). There were no significant differences between groups in the overall amount of spray used ($t(28) = 1.47$, $p = .15$; OT: 14.58 ml ± 3.85; PL: 16.21 ml ± 1.2). Secondary analyses are reported to explore whether treatment induced changes in amygdala connectivity were impacted by medication adherence (see **Supplementary Results**).

Side effect screening. After each week of the 4-week intervention, participants were screened for potential adverse events, side effects or changes in mood. Detailed information on the reporting of side-effects and changes in mood is provided in (Bernaerts et al. 2020). In short, only minimal, non-treatment specific side effects were reported. At the end of the trial, participants were asked if they thought they had received OT or PL. The majority of participants thought they had received the PL treatment (77.5%). The proportion of participants that believed they had received the OT-treatment was not significantly larger in the actual OT-group (28.5%), compared to the PL-group (17.6%) ($p= .46$).

Neural assessment

MRI data acquisition. A 3.0 Tesla Philips Achieva Ds MR-scanner with a 32-channel phased-array head-coil was used to acquire anatomical images and 7-min resting-state fMRI scans during which participants were instructed to relax (but not sleep), keep their eyes open while staring at a white cross and think of nothing in particular. Note that since participants were recruited to participate in a larger clinical trial assessing the (neural and behavioral) effects of multiple-dose treatment with OT, the fMRI scanning protocol additionally included two other scan modalities (not part of the current report): (i) task-based fMRI scanning and (ii) diffusion tensor imaging; both performed after acquisition of the resting state scan.

MRI scanning parameters. Anatomical imaging consisted of a high-resolution structural volume acquired using a coronal three-dimensional turbo field echo T1-weighted sequence with the following parameters: 182 contiguous coronal slices covering the whole brain and brainstem, slice thickness = 1.2 mm; repetition time (TR) = 9.4 ms; echo time (TE) = 4.6 ms; matrix size = 208 x 207; field-of-view

(FOV) = 250 x 250 mm; in-plane pixel size = 1.2 x 1.2 mm²; acquisition time = 1 min 43 s. Resting-state fMRI images were acquired using a T2*-weighted gradient-echo echo planar imaging (GE-EPI) sequence with the following parameters: TR = 2500 ms; TE = 30 ms; matrix size = 80 x 78, FOV = 200 x 200 mm; flip angle 90°; slice thickness = 2.7 mm, slice gap = 0.4 mm; axial slices = 45; 162 functional volumes; acquisition time = 7 min. Note that the fMRI scanning protocol additionally included two other scan modalities: (i) task-based fMRI scanning and (ii) diffusion tensor imaging. The analyses of these scan modalities are not part of the current report.

MRI data preprocessing. SPM-12 (Wellcome Department of Cognitive Neurology, London, UK) and the CONN functional connectivity toolbox 16.b (Whitfield-Gabrieli and Nieto-Castanon 2012) were used for image preprocessing and statistical analyses implemented in Matlab R2015b (Mathworks). Resting-state fMRI images were spatially realigned, normalized to the standard EPI-template of the Montreal Neurological Institute (MNI-152) and resampled into 3-mm isotropic voxels. Realignment parameters were modeled as regressors of no-interest and white matter and cerebrospinal fluid were removed as confounds following the implemented CompCor-strategy (Behzadi et al. 2007) in the CONN toolbox. Residual time-series of the resting-state images were then band-pass filtered (0.009 < f < 0.08Hz).

Head motion. Given the potential confounding effects of micro-movements on resting-state functional connectivity (Power et al. 2012; Van Dijk et al. 2012), all reported analyses were performed on 'scrubbed' data (Power et al. 2012), i.e., censoring frames displaying frame-wise displacement (FD) exceeding > 0.5 mm or frame-wise changes in brain image intensity exceeding >0.5 Δ%BOLD. One subject with a mean FD of > 0.5 (requiring the removal of 98.5% of images) was excluded from the analysis (see Consort diagram in **Figure 1**). On average, approximately 19.2% of the frames (1 min 17 sec out of 7 min) were removed per subject. As visualized in **Supplementary Figure 1**, no group- (OT, PL) or session-related (T0, T1, T2, T3) differences were revealed in mean FD or the percentage of scrubbed frames. Additionally, in **Supplementary Results**, secondary analyses are reported with mean FD added as a 'nuisance' covariate in the main mixed-effects analyses. This practice has been shown to provide substantial additional cleansing of motion-related effects (see (Fair et al. 2012). Also secondary analyses, further excluding subjects with high motion (> 50 % of scrubbed frames) are reported.

Supplementary Results

Secondary analysis

Head motion. Secondary analyses were performed including mean frame-wise displacement (mean FD) as a 'nuisance' covariate into the mixed-effect analysis to determine the extent to which our results pattern was robust to correction for inter-individual differences in head motion. Overall, the pattern of obtained results was consistent with our main findings; indicating a significant main effect of treatment ($F(1, 36)= 6.21$; $p < .017$; $\eta^2 = .15$) and treatment-by-connection interaction ($F(2, 2003)= 3.88$; $p < .001$; $\eta^2 = .017$). Also re-exploration of the mixed model with exclusion of subjects/ sessions with high motion (> 50 % scrubbed frames) (T1 excluded for 2 OT; T2 excluded for 2 OT; T3 excluded

for 2 OT, 1 PL) yielded a pattern of results qualitatively similar to the primary analysis (treatment effect: $F(1, 34.56) = 4.14$; $p < .049$; $\eta^2 = .11$) (treatment-by-connection interaction: $F(1, 1864) = 3.34$; $p < .001$; $\eta^2 = .016$).

Medication adherence. Pearson correlation analyses were performed to explore potential relationships between nasal spray usage and treatment-induced changes in amygdala-OFC connectivity. Overall, changes in amygdala-OFC connectivity (mean across sessions T1, T2 and T3) were not found to be significantly associated with the amount of spray used, either in the IN-OT ($r = .26$; $p = .30$) or PL group ($r = .20$; $p = .52$).

Supplementary Table 1

Detailed information on comorbidities and medication use for participants of the oxytocin and placebo treatment groups.

Comorbidities were screened through self-report (with the explicit mentioning of examples in the screening interview including e.g., ADHD, depression, dyscalculia, dyslexia). Current psychoactive medication use was defined as use within three months before study enrollment. None of the participants reported a change in medication regime between the baseline (T0) and four-week follow-up session (T2). Two participants reported a change in medication use at the one-year follow-up session (T3).

		Comorbidities	Medication use (T0 till T2)	Change Medication use at T3
Oxytocin group		**N= 7**	**N= 5**	**N=2**
	i	ADHD	Abilify, Tegretol	
	ii	Depression	Welbutrine XR, Leviron, Cymbalta	Cymbalta, Trazolan, Edronax, Depakine
	iii	Depression, ADD	Trazodone Mylan, Medikinet	
	iv	Bipolar disorder	Maniprex, Bellozal, Mometasone	
	v	ADHD,Dyslexia	/	
	vi	ADHD, Depression	/	
	vii	Dyslexia	/	
	viii	/	Risperdal, Venlafaxine	
	ix	/	/	Sertraline
Placebo group		**N= 2**	**N= 2**	
	i	ADHD	/	
	ii	ADHD	/	
	iii	/	Zolpidem, Remergon, Rilatine	
	iv	/	Trazodone, Escitalopram	

ADHD: attention deficit hyperactivity disorder; ADD: attention deficit disorder

Supplementary Table 2

Treatment-induced changes in behavior.

As reported in more detail in Bernaerts et al. (2020) (Bernaerts et al. 2020), we previously explored behavioral improvements as a result of the four-week IN-OT treatment in the same patient sample in terms of social functioning (Social Responsiveness Scale - Adult version: SRS-A), repetitive behaviors (Repetitive Behavior Scale – Revised: RBS-R) and attachment avoidance (State Adult Attachment Measure: SAAM).

- The SRS-A (self-report) (64 items) (Constantino et al. 2003) comprises four subscales examining social communication, social awareness, social motivation and rigidity/repetitiveness, using a four-point Likert-scale. SRS-A raw total scores were adopted.
- The RBS-R (self-report) (43 items) (Lam and Aman 2007) examines a heterogeneous set of repetitive behaviors including stereotypic behavior, self-injurious behavior, compulsive behavior, ritualistic behavior, sameness behavior and restricted interests behavior, using a four-point Likert-scale.
- The SAAM (self-report) (Gillath et al. 2009) comprises three subscales, of which one subscale assesses attachment avoidance (e.g., “If someone tried to get close to me, I would try to keep my distance”) (7 items) using a seven-point Likert-scale.

In short, behavioral improvements were evident immediately after treatment (T1) and until four weeks (T2) and one year (T3) post-treatment in repetitive behaviors (RBS-R) and feelings of avoidant attachment (SAAM). While the oxytocin group also reported improvements in social symptoms (SRS-A), these improvements were not treatment-specific (i.e., comparable improvements were evident in the placebo group).

The table below lists for each questionnaire the mean pre-to-post change scores separately for each treatment group (oxytocin, placebo) and assessment session (T1, T2, T3). *T*- and *p*-values correspond to single-sample *t*-tests assessing within-group changes from baseline separately for the oxytocin and placebo group. Cohen’s *d* effect sizes of between-group differences (change from baseline_{OT}–change from baseline_{PL})/pooled SD) are reported where 0.2 is indicative of a small effect, 0.5 a medium effect and 0.8 a large effect.

	Oxytocin				Placebo				Between-group difference
	N	**Mean ± SD**	***T*-value**	***p***	**N**	**Mean ± SD**	***T*-value**	***p***	**Cohen's *d***
Multiple-dose effect (T1)									
SRS-A	22	-5.55 ± 11.40	-2.28	0.033	18	-1.06 ± 10.01	-0.45	0.66	-0.42
RBS-R	22	-4.77 ± 6.47	-3.46	0.002	17	-1.76 ± 4.75	-1.53	0.15	**-0.63**
SAAM avoidance	22	-0.40 ± 0.71	-2.63	0.016	18	0.06 ± 0.98	0.24	0.81	**-0.61**
Four-week follow-up (T2)									
SRS-A	22	-5.64 ± 12.57	-2.10	0.048	18	-7.67 ± 12.09	-2.69	0.015	0.22
RBS-R	22	-4.91 ± 6.33	-3.64	0.002	17	-2.35 ± 3.43	-2.83	0.012	**-0.50**
SAAM avoidance	22	-0.38 ± 0.70	-2.58	0.018	18	-0.06 ± 0.76	-0.35	0.73	**-0.53**
One-year follow-up (T3)									
SRS-A	22	-8.59 ± 20.95	-1.92	0.07	18	-6.72 ± 21.01	-1.36	0.19	-0.12
RBS-R	22	-4.91 ± 9.46	-2.43	0.02	17	-0.41 ± 4.27	-0.40	0.70	**-0.98**
SAAM avoidance	22	-0.52 ± 1.18	-2.07	0.05	18	0.0 ± 0.75	0.00	1.00	**-0.80**

SRS-A = Social Responsiveness Scale adult version, RBS-R = Repetitive Behavior Scale – Revised, SAAM = State Adult Attachment Measure, Negative scores indicate pre-to-post improvement.

T- and *p*-values correspond to single-sample *t*-tests assessing within-group changes from baseline separately for the oxytocin and placebo group. Cohen's *d* effect sizes of between-group differences (change from $baseline_{OT}$–change from $baseline_{PL}$)/pooled SD) are reported where 0.2 is indicative of a small effect, 0.5 a medium effect and 0.8 a large effect. Data printed in bold show Cohen's *d* effect sizes equal to or larger than .50 (medium-sized effect).

Supplementary Table 3

Exploratory whole-brain analyses were performed by assessing treatment-induced changes in functional connectivity between bilateral amygdala (seeds) and all other regions of the cortical (n=91) and subcortical (n = 13) Harvard-Oxford atlas, as well as the cerebellar parcellation of the AAL Atlas (n = 26).

An overall pattern of reduced amygdala connectivity in the oxytocin group, compared to the placebo group was identified. Connections displaying significant treatment effects (independent *t*-test, two-sided) are reported separately for each assessment session (T1, T2, T3) at an uncorrected p<.05 threshold. Negative *t*-values indicate oxytocin < placebo.

Session	Seed	ROI		*t*	*p*
T1					
	Right Amygdala	Parietal Operculum Cortex	Right	-3.35	0.002
		Central Opercular Cortex	Left	-2.81	0.008
		Supramarginal Gyrus, anterior	Right	-2.78	0.009
		Cerebellum 3	Left	2.38	0.023
		Vermis		2.16	0.038
		Putamen	Right	-2.08	0.044
		Cerebellum 6	Left	-2.08	0.045
	Left Amygdala	Orbitofrontal Cortex	Right	-2.47	0.019
		Cerebellum 3	Left	2.45	0.019
		Putamen	Right	-2.33	0.026
		Temporal Pole	Left	-2.25	0.031
		Central Opercular Cortex	Left	-2.16	0.038
T2					
	Right amygdala	Orbitofrontal Cortex	Left	-3.13	0.004
		Orbitofrontal Cortex	Right	-3.05	0.004
		Temporal Pole	Right	-2.90	0.006
		Middle Temporal Gyrus, posterior	Right	-2.76	0.009
		Temporal Fusiform Cortex, anterior	Right	-2.61	0.013
		Vermis		-2.54	0.016
		Supramarginal Gyrus, posterior	Left	-2.47	0.019
		Brainstem		-2.42	0.021
		Temporal Pole	Left	-2.41	0.021
		Superior Temporal Gyrus, anterior	Right	-2.35	0.025
		Temporal Fusiform Cortex, anterior	Left	-2.19	0.035

	Left Amygdala	Temporal Fusiform Cortex, anterior	Right	-3.27	0.002
		Temporal Pole	Left	-2.74	0.010
		Orbitofrontal Cortex	Left	-2.50	0.017
		Temporal Pole	Right	-2.43	0.020
		Supramarginal Gyrus, posterior	Left	-2.30	0.027
		Cerebellum 6	Left	-2.29	0.028
		Inferior Temporal Gyrus	Left	-2.26	0.030
		Cerebellum 8	Left	-2.20	0.034
		Brainstem		-2.15	0.039
		Temporal Fusiform Cortex, anterior	Left	-2.12	0.041
T3					
	Right Amygdala	Cerebellum 7	Right	-2.46	0.020
		Putamen	Right	-2.26	0.032
		Caudate	Right	-2.23	0.033
	Left Amygdala	Cerebellum 7	Rigth	-3.09	0.004
		Orbitofrontal Cortex	Right	-3.02	0.005
		Caudate	Right	-2.99	0.006
		Putamen	Right	-2.91	0.007
		Orbitofrontal Cortex	Left	-2.52	0.017
		Superior Temporal Gyrus, posterior	Right	-2.19	0.037
		Middle Temporal Gyrus, anterior	Left	-2.17	0.038
		Globus Pallidum	Right	-2.07	0.047

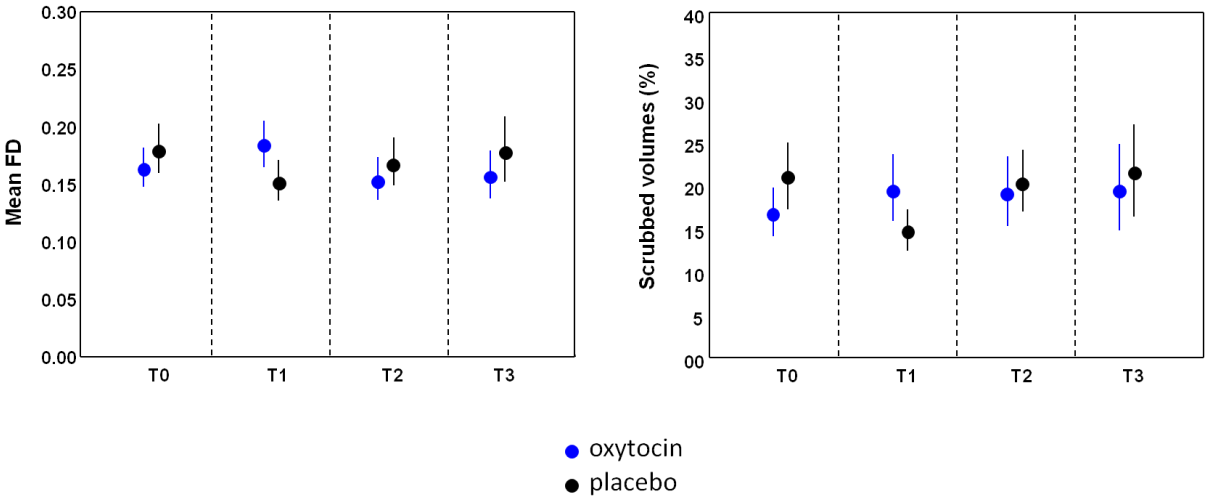
T1: assessment session immediately after the four-week treatment (at least 24 hours after the final administration); T2: assessment session four-weeks post-treatment; T3: assessment session one-year post-treatment. ROI: Region of interest.

Supplementary Figure 1

Head motion analysis of the resting-state fMRI scans.

For all participants, head motion (mean frame-wise displacement (mean FD)) of the resting-state fMRI scans was assessed at baseline (**T0**); after the four-week treatment (**T1**); and at the two follow-up sessions, four-weeks (**T2**), and one-year post-treatment (**T3**).

Mean FD scores were not significantly different between the oxytocin and placebo groups (all, $t < 1.22$; $p > .22$). Also no significant effects of session were revealed in the oxytocin ($F(3, 68.20) = 1.94$, $p = .13$) or placebo group ($F(3, 56.47) = 1.32$, $p = .28$). Also for the percentage of scrubbed frames, no significant group- (all, $t < 1.02$; $p > .31$) or session-related differences were revealed (oxytocin: $p > .82$; placebo: $p > .34$).



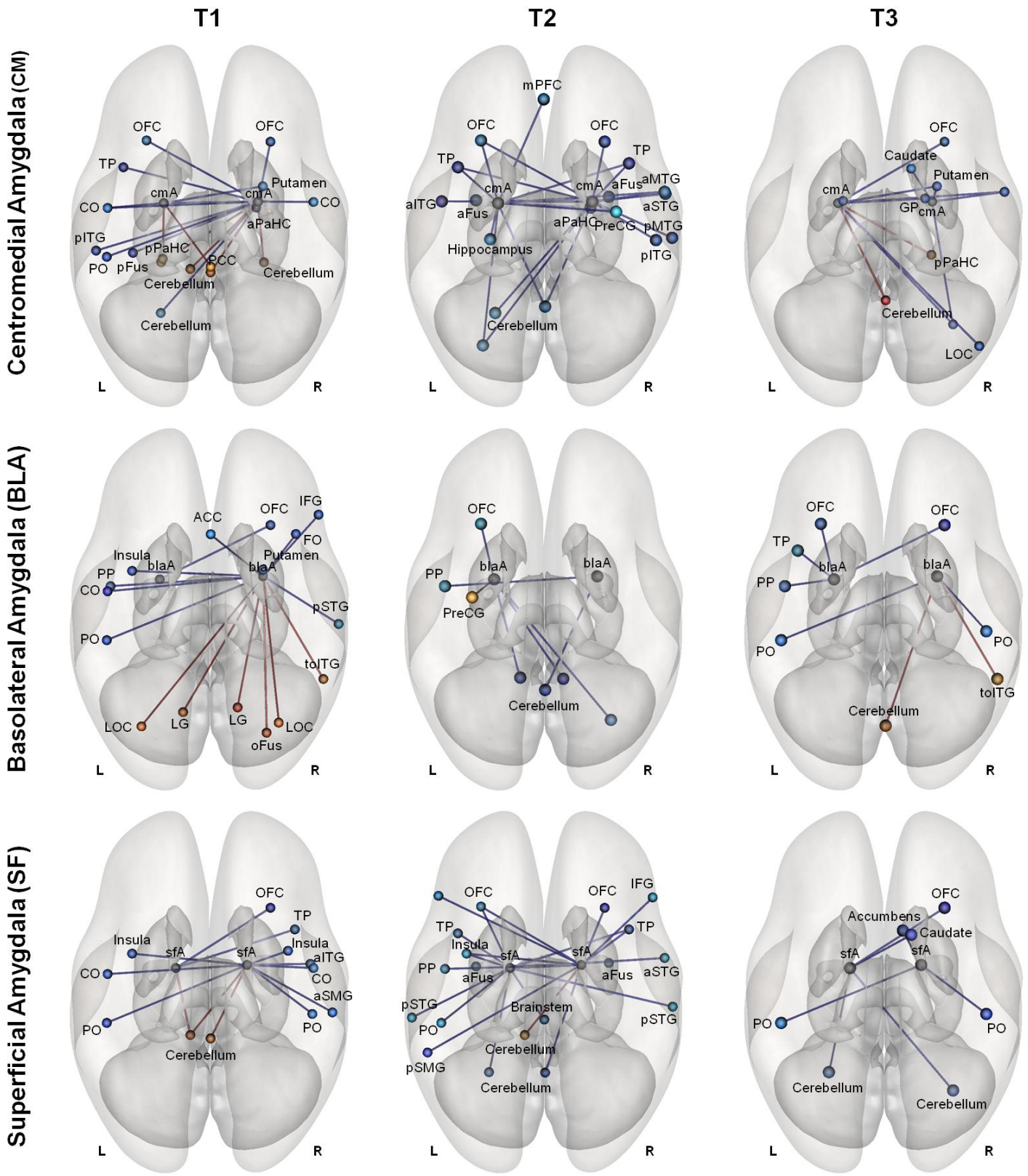
Supplementary Figure 2

Whole-brain connectivity analyses of amygdala sub-regions.

Regions-of-interest of distinct amygdala sub-regions covering the centromedial (CM), basolateral (BLA) and superficial (SF) amygdala were defined using probabilistic maps from the SPM anatomy Toolbox (Eickhoff et al. 2005) (as previously adopted in (Eckstein et al. 2017)) to explore treatment-induced changes in sub-regional amygdala connectivity to other regions of the cortical (n = 91), and subcortical (n = 13) Harvard-Oxford Atlas, and the cerebellar parcellation of the AAL Atlas (n = 26).

Similar to the main analysis reporting the effect of treatment on (whole) amygdala connectivity (**Figure 3**), the sub-regional connectivity analysis also revealed an overall pattern of *reduced* amygdala connectivity in the oxytocin group, compared to the placebo group. Connections displaying significant treatment effects (blue connections: oxytocin < placebo; red connections: oxytocin > placebo) are reported separately for each assessment session (T1, T2, T3) at an uncorrected p <.05 threshold (two-sided). T1: assessment session immediately after the four-week treatment (at least 24 hours after the final administration); T2: assessment session four weeks post-treatment; T3: assessment session one-year post-treatment.

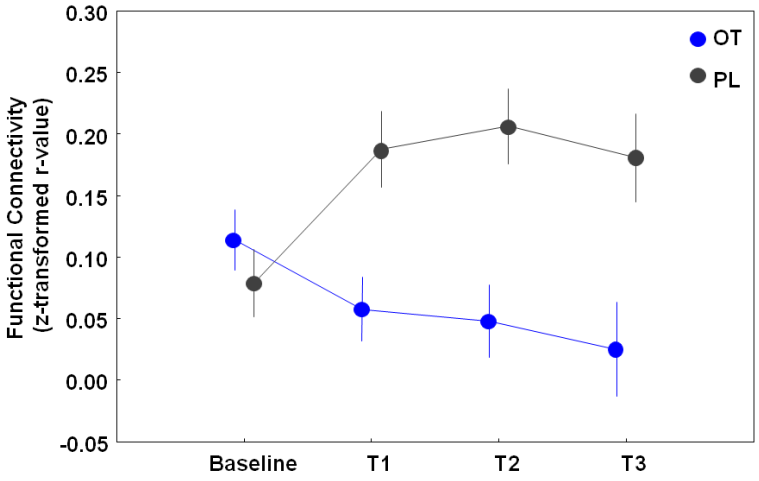
cmA: Centromedial Amygdala; blaA: Basolateral Amygdala; sfA: Superficial Amygdala; PO: Parietal Operculum Cortex; CO: Central Opercular Cortex; OFC: Orbitofrontal Cortex; mPFC: medial Prefrontal Cortex; TP: Temporal Pole; aMTG: anterior Middle Temporal Gyrus; pMTG: posterior Middle Temporal Gyrus; aSTG: anterior Superior Temporal Gyrus; pSTG: posterior Superior Temporal Gyrus; pITG: aITG: anterior Inferior Temporal Gyrus; posterior Inferior Temporal Gyrus; toITG: temporal-occipital Inferior Temporal Gyrus; aFus: anterior Fusiform Cortex; pFus: posterior Fusiform Cortex; oFus: occipital Fusiform Cortex; aPaHC: anterior Parahippocampal cortex; pPaHC: posterior Parahippocampal cortex; aSMG: anterior Supramarginal Gyrus; pSMG: posterior Supramarginal Gyrus; PreCG: precentral Gyrus; LOC: Lateral Occipital Cortex; LG: Lingual Gyrus; PP: Planum Polare; FO: Frontal Operculum; ACC: anterior Cingulate Cortex; IFG: Inferior Frontal Cortex; GP: Globus Pallidum; L: left; R: right.


T1
T2
T3
Centromedial Amygdala (CM)
Basolateral Amygdala (BLA)
Superficial Amygdala (SF)
OFC
TP
Putamen
cmA
CO
aPaHC
pITG
pPaHC
PO
pFus
PCC
Cerebellum
mPFC
aITG
aFus
aMTG
aSTG
pMTG
PreCG
Hippocampus
Caudate
GP
LOC
ACC
IFG
FO
Insula
blaA
PP
pSTG
tolTG
LG
oFus
sfA
aSMG
Accumbens
Brainstem
pSMG
L
R

Supplementary Figure 2

Visualization of raw connectivity scores.

Mean amygdala-orbitofrontal connectivity scores (z-transformed r-values) are visualized separately for the oxytocin (OT) and placebo (PL) group at baseline, after the four-week (oxytocin/ placebo) treatment (T1); and at the two follow-up sessions, four weeks (T2) and one year after cessation of the treatment (T3). As visualized, the IN-OT treatment primarily induced a reduction in the strength of coupling, rather than affecting the directionality of coupling (i.e. to anti-correlations; negative z-transformed r-values).



References

1. Behzadi Y, Restom K, Liau J, Liu TT (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage* **37**: 90-101.

2. Bernaerts S, Boets B, Bosmans G, Steyaert J, Alaerts K (2020). Behavioral effects of multiple-dose oxytocin treatment in autism: A randomized, placebo-controlled trial with long-term follow-up. *Molecular Autism* **11**.

3. Constantino JN, Davis SA, Todd RD, Schindler MK, Gross MM, Brophy SL*, et al.* (2003). Validation of a brief quantitative measure of autistic traits: Comparison of the social responsiveness scale with the autism diagnostic interview-revised. *Journal of Autism and Developmental Disorders* **33**: 427-433.

4. Eckstein M, Markett S, Kendrick KM, Ditzen B, Liu F, Hurlemann R*, et al.* (2017). Oxytocin differentially alters resting state functional connectivity between amygdala subregions and emotional control networks: Inverse correlation with depressive traits. *Neuroimage* **149**: 458-467.

5. Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K*, et al.* (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* **25**: 1325-1335.

6. Fair DA, Nigg JT, Iyer S, Bathula D, Mills KL, Dosenbach NU*, et al.* (2012). Distinct neural signatures detected for ADHD subtypes after controlling for micro-movements in resting state functional connectivity MRI data. *Front Syst Neurosci* **6**: 80.

7. Gillath O, Hart J, Noftle EE, Stockdale GD (2009). Development and validation of a state adult attachment measure (SAAM). *Journal of Research in Personality* **43**: 362-373.

8. Guastella AJ, Hickie IB, McGuinness MM, Otis M, Woods EA, Disinger HM*, et al.* (2013). Recommendations for the standardisation of oxytocin nasal administration and guidelines for its reporting in human research. *Psychoneuroendocrinology* **38**: 612-625.

9. Lam KS, Aman MG (2007). The Repetitive Behavior Scale-Revised: independent validation in individuals with autism spectrum disorders. *J Autism Dev Disord* **37**: 855-866.

10. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* **59**: 2142-2154.

11. Van Dijk KR, Sabuncu MR, Buckner RL (2012). The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage* **59**: 431-438.

12. Whitfield-Gabrieli S, Nieto-Castanon A (2012). Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect* **2**: 125-141.