

Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans

René Hurlemann, Alexandra Patin, Oezguer A. Onur, Michael X. Cohen, Tobias Baumgartner, Sarah Metzler, Isabel Dziobek, Jürgen Gallinat, Michael Wagner, Wolfgang Maier, and Keith M. Kendrick

Supplemental Information section

Experiment 1

Neuropsychological screening for oxytocin- and placebo-treated male subjects

A total of 48 healthy male volunteers (n = 24 placebo, mean \pm SD age 25.2 \pm 2.5 years; n = 24 oxytocin, age 26.7 \pm 2.2 years) were tested with the LPS-4 ('Leistungsprüfsystem' subtest 4) to assess non-verbal reasoning IQ (Horn, 1983), the DST (digit span test implemented in the HAWIE-R and WAIS-R) (Wechsler, 1981; Tewes, 1991) and the RAVLT (Rey Auditory Verbal Learning Test) (Rey, 1941, 1964; Helmstaedter et al., 1981) to assess immediate verbal memory span, new learning, susceptibility to interference, and delayed recall, and the FEEST (Facial Expressions of Emotion: Stimuli and Tests) (Young et al., 2002) to assess facial emotion recognition skills.

Table S1. Neuropsychological tests: Oxytocin- vs. placebo-treated subjects

Characteristic	Placebo (n = 24) Raw score(SD)	Oxytocin (n = 24) Raw score(SD)	<i>p</i> value
RAVLT			
Trial 1-5 ^a	55.50(8.63)	56.88(8.78)	0.587
Trial 5 ^b	13.29(1.68)	13.79(1.22)	0.244
Trial 6 Retention ^b	12.29(2.26)	13.00(1.84)	0.239
Trial 7 Delayed recall ^b	12.25(2.64)	12.42(2.36)	0.819
DST ^c	17.21(3.51)	16.42(3.34)	0.427
LPS-4 ^d	32.96(3.50)	31.71(3.88)	0.247
FEEST			
Happy ^e	7.88(0.34)	7.88(0.34)	1.000
Sad ^e	5.79(1.14)	5.75(1.11)	0.899
Disgust ^e	4.17(1.71)	4.42(1.53)	0.596
Fear ^e	5.29(1.49)	5.58(1.64)	0.522
Anger ^e	6.58(1.02)	6.67(0.82)	0.756
Surprise ^e	6.63(1.13)	6.83(1.09)	0.520

RAVLT, Rey Auditory Verbal Learning Test; DST, Digit Span Test; LPS-4, a non-verbal reasoning IQ test; FEEST, Facial expressions of emotion: stimuli and tests.

^a Maximum possible score 75.

^b Maximum possible score 15.

^c Maximum possible score of sum 28.

^d Maximum possible score 40, raw score 32 is IQ 115-120.

^e Maximum possible score 9.

Experiment 3

Neuropsychological assessment of Urbach-Wiethe patients A.M. and B.G.

As outlined in the main article, we used an extensive neuropsychological battery to test the patients for memory, attention, IQ, and visuospatial construction skills, as well as screen them for psychopathologies. The RAVLT (Rey Auditory Verbal Learning Test) (Rey, 1941, 1964; Helmstaedter et al., 1981) was used to assess immediate verbal memory span, new learning, susceptibility to interference, and delayed recall. The CFT (Rey-Osterrieth Complex Figure Test) was included to test incidental visual memory as well as the visuospatial constructional ability (Rey, 1941; Osterrieth, 1944). To examine figural learning and memory, we used the DCS ('Diagnostikum für Cerebralschädigung') developed for brain-damaged patients (Weidlich and Lamberti, 2001). Furthermore, speed of attention and mental flexibility were assessed with the TMT (Trail Making Test) (Reitan, 1955), and short-term concentration ability was tested with the non-verbal d2 test ('Aufmerksamkeits- und Belastungstest d2') (Brickenkamp, 1995). Color-word interference was tested with the Stroop test (Stroop, 1935; Baeumler, 1985). To assess perseveration and abstract reasoning, we included the WCST-64, a normed modification of the Wisconsin Card Sorting Test that utilizes one deck of cards (Kongs et al., 2000). To assess speed and flexibility of verbal thought processes, we included the RWT ('Regensburger Wortflüssigkeitstest'), a verbal fluency test (Aschenbrenner et al., 2000). The HAWIE-R ('Hamburg-Wechsler Intelligenztest für Erwachsene in revidierter Fassung'), a German-language adaptation of the WAIS-R (Wechsler Intelligence Test for Adults - Revised) (Wechsler, 1981; Tewes, 1991) provided a measure of verbal, performance, and full-scale IQ; the Mehrfachwahl-Wortschatztest provided a verbal IQ (Lehrl, 1978). The BDI-II (Beck Depression Inventory-Second Edition) (Beck et al., 1996), the HDRS-21 (Hamilton Depression Rating Scale) (Hamilton, 1960), and the HARS (Hamilton Anxiety Rating Scale) (Hamilton, 1959) were used for depression and anxiety screening. For further assessment of psychological and psychopathological symptom load, we included the SCL-90-R (Symptom Checklist-90-Revised) instrument (Derogatis, 1993; Franke, 1995). Perception and understanding of non-verbal communicative signals of emotion (facial expressions/voice) were assessed with the TAB ('Tübinger Affekt Batterie'), a German-language version of the FAB (Florida Affect Battery) (Bowers et al., 1991; Breitenstein et al., 1996).

Table S2. Neuropsychological tests: Urbach-Wiethe patients

Characteristic	A.M.	B.G.
RAVLT (percentile)		
Trial 1-5	40	80
Trial 5	45-55	60-75
Trial 6 Retention	25-30	35-45
Trial 7 Delayed recall	15-20	25-30
DCS (percentile)		
Trial 1	15	34.5
Trial 2	< 6	34
Trial 3	2.2	26.8
Trial 4	< 4.3	25.6
Trial 5	< 8.6	22.5-30
Trial 6	< 4.3	22.5
Lability Index (over 6 trials; raw)	0	0
d2 (percentile)		
Total correct	8	7
TMT (percentiles)		
Part A	76	58
Part B	40	46
Stroop (T scores)		
Color-word list	66	79
Color-box naming	66	62
Interference	73	62
WCST-64		
Correct (raw score) ^a	45	55
Errors (percentile)	18	58
Conceptual level responses (percentage)	14	61
RWT (percentile)		
Phonetic fluency (P-words; 2 min)	< 2	< 2
Semantic fluency (animals; 1 min)	23	15
CFT (raw)		
Copy ^b	34	34
Recall ^c	7	16

RAVLT, Rey Auditory Verbal Learning Test; CFT, Complex Figure Test; DCS, a figural learning and memory test; d2, a short-term concentration performance test; TMT, Trail Making Test; Stroop, Stroop test; WCST-64, Wisconsin Card Sorting Test – 64; RWT, a verbal fluency test.

^a Maximum possible score 64.

^b Maximum possible score 36.

^c Recall 30 min after copy, maximum possible score 36.

Table S3. IQ tests: Urbach-Wiethe patients

Test	A.M.	B.G.
HAWIE-R		
Verbal IQ (estimated IQ score)	99	94
General knowledge ^a	9	9
Repeating numbers ^a	13	12
Vocabulary ^a	9	10
Arithmetic thinking ^a	10	8
General comprehension ^a	13	10
Finding similarities ^a	8	9
Performance IQ	103	97
Complementing pictures ^a	12	8
Organizing pictures ^a	10	11
Mosaic test ^a	9	9
Arranging figures ^a	9	6
Number-symbol test ^a	9	11
Full-scale IQ	101	96
LPS-4 (estimated IQ)	107.5	92.5
MWT-B (estimated IQ)	100	106

HAWIE-R, German-language adaptation of the WAIS-R; LPS-4, a non-verbal reasoning IQ test; MWT-B, a verbal IQ test based on lexical decisions.

^a Value points; maximum possible score 19.

Table S4. Psychopathological symptoms: Urbach-Wiethe patients

Questionnaire	A.M.	B.G.
BDI-II (raw) ^a	5	5
HDRS-21 (raw) ^b	0	0
HARS (raw) ^c	0	1
SCL-R-90 (T scores)		
Somatization	37	29
Obsessive-compulsive	55	53
Interpersonal sensitivity	45	54
Depression	38	59
Anxiety	43	36
Anger-hostility	52	52
Phobic anxiety	43	43
Paranoid ideation	38	46
Psychoticism	39	39
Global severity index	40	48

BDI-II, Beck Depression Inventory; HDRS-21, Hamilton Depression Rating Scale-21; HARS, Hamilton Anxiety Rating Scale; SCL-R-90, Symptom Checklist-90-Revised.

^a Maximum score 63, cutoff score for 'not depressed' 9

^b Maximum score 70, cutoff score for 'not depressed' 15

^c Maximum score 56, cutoff score for 'mildly anxious' 17

Table S5. Performance of Urbach-Wiethe patients on the Florida Affect Battery (FAB)

Subtest	A.M.	Z score	B.G.	Z score	Controls (n = 10)
Facial identity discrimination ^a	91.67	-	100	-	100(0)
Facial affect discrimination ^b	64.28	-3.98	71.43	-2.90	90.60(6.61)
Facial affect naming ^c	100	0.81	86.67	-0.54	92.00(9.84)
Facial affect selection ^d	100	1.01	93.33	-0.25	94.67(5.26)
Facial affect matching ^e	86.67	-0.54	93.33	0.14	92.00(9.84)
Nonemotional prosody discrimination ^f	100	0.44	100	0.45	97.27(6.14)
Emotional prosody discrimination ^g	100	0.62	92.3	-1.45	97.69(3.72)
Name the emotional prosody ^h	93.33	0.50	73.3	-0.75	85.33(15.96)
Conflicting emotional prosody, congruent ⁱ	62.5	-6.64	100	-1.90	97.50(5.27)
Conflicting emotional prosody, incongruent ⁱ	79.17	-1.36	96.67	-0.04	92.08(9.51)
Match emotional prosody to an emotional face ^j	86.67	-0.20	93.33	0.47	88.67(9.96)
Match emotional face to the emotional prosody ^k	100	0.99	100	0.99	92.28(7.79)

^a Tests the ability to distinguish between identical and different faces; ^b Tests the ability to distinguish between identical and different affects on faces; ^c Tests the ability to name an affect; ^d Tests the ability to select a specific affect named by the examiner; ^e Tests the ability to match one face showing a specific affect with a different face showing the same affect; ^f Tests the ability to distinguish between a question and a statement in prosody; ^g Tests the ability to distinguish between identical and different affects in prosody; ^h Tests the ability to name a specific affect heard in a sentence; ⁱ Tests the ability to name a specific affect heard in a sentence, whether the content is congruent or incongruent to the affect; ^j Tests the ability to match a specific affect heard in a sentence to a face showing the same affect; ^k Tests the ability to match an affect on a face to the same affect heard in a sentence.

References

- Aschenbrenner S, Tucha O, Lange KW (2000) Regensburger Wortflüssigkeitstest. Göttingen, Germany: Hogrefe.
- Baeumler G (1985) FWIT: Farbe-Wort-Interferenztest [The Stroop-Test] Göttingen, Germany: Hogrefe.
- Beck AT, Steer RA, Brown GK (1996) Manual for the Beck Depression Inventory-II. San Antonio: Psychological Corporation.
- Bowers D, Blonder LX, Heilman KM (1991) Florida Affect Battery. Gainesville, FL: University of Florida.
- Breitenstein C, Daum I, Ackermann H, Lütgehetmann R, Müller E (1996) Erfassung der Emotionswahrnehmung bei zentralnervösen Läsionen und Erkrankungen: Psychometrische Gütekriterien der "Tübinger Affekt Batterie". *Neurol Rehabil* 2:93–101.
- Brickenkamp R (1995) Aufmerksamkeitsbelastungstest 'd2', erweiterte und neu gestaltete Auflage. *Diagnostica* 41:291–296.
- Derogatis LR (1993) The Brief Symptom Inventory: Administration, Scoring and Procedures. Minneapolis, MN: National Computer Systems.
- Franke GH (1995) SCL-90-R. Die Symptom Checkliste von Derogatis, Deutsche version. Göttingen, Germany: Beltz Test.
- Hamilton M (1959) The assessment of anxiety states by rating. *Br J Med Psychol* 32:55–55.
- Hamilton M (1960) A rating scale for depression. *J Neurol Neurosurg Psychiatry* 23:56–62.
- Helmstaedter C, Lendt M, Lux S (1981) VLMT Verbaler Lern und Merkfähigkeitstest. Göttingen, Germany: Beltz Test.
- Horn, W (1983) L-P-S Leistungsprüfsystem. Göttingen, Germany: Hogrefe.
- Kongs SK, Thompson LL, Iverson GL, Heaton RK (2000) The Wisconsin Card Sorting Test (WCST-64): Computer Version 2 Research Edition. Odessa, FL: Psychological Assessment Resources.
- Lehrl S (1978) Mehrfachwahl-Wortschatz-Intelligenztest MWT-B. Erlangen, Germany: Verlag Dr. med. Straube.
- Osterrieth PA (1944) Le test de copie d'une figure complexe. *Arch Psychol* 30:206–356.
- Reitan RM (1955) The relation of the trail-making test to organic brain damage. *J Consult Psychology* 19:393–394.
- Rey A (1941) L'examen psychologique dans les cas d'encephalopathie traumatique. *Arch Psychol* 28:286–340.
- Rey A (1964). L'examen clinique en psychologie. Paris, France: Presses Universitaires de France.
- Stroop JR (1935) Studies of interference in serial verbal reactions. *J Exp Psychol* 18:643–662.
- Tewes U (1991) Hamburg-Wechsler-Intelligenztest für Erwachsene, Revision. Bern, Switzerland: Hans Huber.
- Young AW, Perrett DI, Calder AJ, Sprengelmeyer R, Ekman P (2002) Facial expressions of emotion: stimuli and tests (FEEST). Bury St. Edmunds: Thames Valley Test Company.
- Wechsler D (1981) Wechsler adult intelligence scale-revised. New York: The Psychological Corporation.
- Weidlich S, Lamberti G (2001) DCS – Diagnosticum für Cerebralschädigung (DCS). Ein visueller Lern- und Gedächtnistest nach F. Hiller. Bern, Switzerland: Hans Huber.