Oxytocin facilitates protective responses to aversive social stimuli in males

Nadine Striepens[a,1], Dirk Scheele[a,1], Keith M. Kendrick[b,1], Benjamin Becker[a,1], Lea Schäfer[a], Knut Schwalba[a], Jürgen Reul[c], Wolfgang Maier[a,d], and René Hurlemann[a,2]

[a]Department of Psychiatry, University of Bonn, 53105 Bonn, Germany; [b]Key Laboratory for Neuroinformation, School of Life Science and Technology, University of Electronic Science and Technology of China, 610054 Chengdu, People's Republic of China; [c]Beta Clinic, 53227 Bonn, Germany; and [d]German Center for Neurodegenerative Diseases, 53175 Bonn, Germany

Edited by Bruce S. McEwen, The Rockefeller University, New York, NY, and approved September 18, 2012 (received for review May 24, 2012)

The neuropeptide oxytocin (OXT) can enhance the impact of positive social cues but may reduce that of negative ones by inhibiting amygdala activation, although it is unclear whether the latter causes blunted emotional and mnemonic responses. In two independent double-blind placebo-controlled experiments, each involving over 70 healthy male subjects, we investigated whether OXT affects modulation of startle reactivity by aversive social stimuli as well as subsequent memory for them. Intranasal OXT potentiated acoustic startle responses to negative stimuli, without affecting behavioral valence or arousal judgments, and biased subsequent memory toward negative rather than neutral items. A functional MRI analysis of this mnemonic effect revealed that, whereas OXT inhibited amygdala responses to negative stimuli, it facilitated left insula responses for subsequently remembered items and increased functional coupling between the left amygdala, left anterior insula, and left inferior frontal gyrus. Our results therefore show that OXT can potentiate the protective and mnemonic impact of aversive social information despite reducing amygdala activity, and suggest that the insula may play a role in emotional modulation of memory.

emotion | functional imaging | psychophysiology | cognition

Current concepts of the neuromodulatory role of the peptide hormone oxytocin (OXT) in human cognition and behavior emphasize its prosocial effects. Ample evidence for this view comes from a plethora of behavioral experiments in healthy subjects (1), which have demonstrated beneficial effects of a single intranasal dose (24–48 IU) of OXT across a wide range of cognitive tasks, probing, for example, interpersonal trust and cooperation (2, 3; see also ref. 4), generosity (5), social recognition (6–8) and related memory (9–11), social reinforcement learning and emotional empathy (12), and social judgments (13–15).

This prosocial perspective on OXT is challenged, however, by evidence that OXT also enhances envy and schadenfreude (gloating) (16), ethno-centrism (including prejudice, xenophobia, and racial bias) (4), and outgroup derogation (17). Moreover, OXT hinders trust and cooperation when social information about interaction partners is lacking (18). Furthermore, OXT appears to negatively bias recollections of maternal care and closeness and to diminish trust and cooperation in insecurely or anxiously attached individuals (19, 20).

In an attempt to reconcile this controversial evidence, it has been proposed that the social effects of OXT could be mediated by reduced anxiety or by an increased perceptual salience of social cues (21). The anxiolytic action of OXT has been confirmed by showing reduced amygdala responses to aversive social stimuli in healthy people (22–25; but see also refs. 26 and 27), and subjects with social phobia (28). It is compatible with decreased endocrine and subjective responses to social stress (29), as well as reduced negative cognitive self-appraisal in individuals scoring high in trait-anxiety (30). In contrast, the social salience hypothesis has gained substantial support from studies demonstrating increased eye contact (31) and improved mind-reading from facial expressions (32) as a result of OXT treatment. Whether these mechanisms quintessentially yield positive or negative social outcomes may vary depending on contextual or person-specific characteristics (21). An alternative view holds that emotional valence may be the key in guiding the social effects of OXT, with it facilitating social approach to positive cues and inhibiting social withdrawal from negative ones (33).

Against this empirical background we performed two experiments, using paradigms in the social-protective domain, to disentangle the valence-specific effects of OXT and to determine if anxiolytic, salience, or approach-enhancing mechanisms are most influential. Specifically, Exp. 1 examined the emotional modulation of the acoustic startle reflex (ASR) and we predicted that anxiolytic effects should lead to overall diminished ASR magnitudes and less pronounced potentiation of the ASR in the context of negative stimuli. On the other hand, the social-salience hypothesis would be compatible with facilitated emotional modulation both with negative and positive stimuli, whereas the social-approach/withdrawal hypothesis would predict differential modulation, with startle responses being reduced to negative (withdrawal) and increased to positive (approach) stimuli. Exp. 2 addressed the question of how OXT shapes the neural correlates of emotion perception and subsequent memory, thus enabling us to characterize the neuromodulatory influence of OXT on information of both immediate and future relevance. Decreased amygdala responses to negative stimuli and reduced memory for them should be a consequence of anxiolytic effects, whereas an increased salience of such stimuli could improve subsequent memory performance. On the other hand, if OXT enhances social approach behavior then one might predict either no effect—because negative stimuli would not normally promote approach behavior—or possibly make them less memorable because of weakening of the normal withdrawal response.

Results

Sample Characteristics. From 80 subjects enrolled in Exp. 1, 11 subjects [OXT, $n = 5$; PLC (placebo: sodium chloride solution), $n = 6$] were excluded from further analysis because of technical failures during data acquisition ($n = 4$), mood alterations ($n = 1$), or lack of distinct startle responses ($n = 6$). Consequently, subsequent analyses included data from 69 subjects (OXT, $n = 36$; PLC, $n = 33$). From the 73 subjects enrolled in Exp. 2, three were excluded from further analysis because of technical failures during data acquisition. Consequently, subsequent analyses were performed on the

Author contributions: N.S., D.S., B.B., and R.H. designed research; N.S., D.S., B.B., L.S., and K.S. performed research; J.R. contributed new reagents/analytic tools; N.S., D.S., K.M.K., B.B., L.S., K.S., and R.H. analyzed data; and N.S., D.S., K.M.K., B.B., W.M., and R.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]N.S., D.S., K.M.K., and B.B. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: renehurlemann@me.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1208852109/-/DCSupplemental.

data obtained from 70 subjects (OXT, $n = 35$; PLC, $n = 35$). In both experiments, treatment groups showed no differences regarding demographics, pretreatment neuropsychological performance, attention, or anxiety (Tables S1 and S2).

Experiment 1. ***Emotion-modulated startle.*** A repeated-measures ANOVA with "group" (PLC, OXT) as between-subjects factor, "valence" (negative, neutral, or positive) as within-subjects factor, and "ASR magnitude" as dependent variable revealed a main effect of valence [$F_{(2, 134)} = 34.49$, $P < 0.01$, $\eta^2 = 0.34$] and an interaction effect of valence and group [$F_{(2, 134)} = 4.09$, $P = 0.02$, $\eta^2 = 0.06$]. Both groups showed a significant linear trend for valence (i.e., negative > neutral > positive), but a comparison of the effect sizes indicated that the emotional modulation in the OXT group [$F_{(1, 35)} = 84.10$, $P < 0.01$, $\eta^2 = 0.71$] was more than twice as large as in the PLC group [$F_{(1, 32)} = 13.78$, $P < 0.01$, $\eta^2 = 0.30$] (Fig. 1*A*). Post hoc Bonferroni-corrected two-sample *t* tests showed that the ASR magnitude in the OXT group was significantly larger than in the PLC group during the viewing of negative [$t_{(67)} = 2.92$, $P = 0.01$, $\eta^2 = 0.11$], but not neutral [$t_{(54.94)} = -1.54$, $P = 0.37$, $\eta^2 = 0.04$] or positive [$t_{(53.91)} = -1.04$, $P = 0.91$, $\eta^2 = 0.02$] pictures. The contrasts "negative minus neutral" [$t_{(57.77)} = 2.56$, $P = 0.01$, $\eta^2 = 0.09$] and "negative minus positive" [$t_{(59.07)} = 2.33$, $P = 0.02$, $\eta^2 = 0.08$] yielded significantly larger difference scores in the OXT group, but there was no OXT effect for the contrast "positive minus neutral" [$t_{(67)} = 0.31$, $P = 0.76$, $\eta^2 < 0.01$] (Fig. 1*B*). In summary, administration of OXT profoundly potentiated the emotional modulation of the ASR specifically for negative stimuli.

Emotion ratings. A repeated-measures ANOVA with "group" (PLC, OXT) as between-subjects factor, "valence" (negative, neutral, positive) as within-subjects factor, and "valence ratings" as dependent variable yielded a main effect of valence [$F_{(1.38, 92.46)} = 270.68$, $P < 0.01$, $\eta^2 = 0.80$], but no interaction effect [$F_{(1.38, 92.46)} = 0.22$, $P = 0.72$, $\eta^2 < 0.01$]. As expected, negative pictures were rated as more aversive than neutral ones and positive pictures as more pleasant than neutral ones (Table S3). For the arousal ratings, there was a main effect of valence [$F_{(2, 134)} = 302.18$, $P < 0.01$, $\eta^2 = 0.82$] but no interaction effect [$F_{(2, 134)} = 0.16$, $P = 0.85$, $\eta^2 < 0.01$]. Negative and positive pictures were therefore rated as equally more arousing than neutral ones in both groups (Table S3).

Baseline startle response. Raw startle response magnitudes from the interstimulus intervals were individually examined. At visual inspection, the overall startle magnitude seemed to be smaller in the OXT group (mean ± SD, 31.0 ± 46.2 μV) than in the PLC group (47.1 ± 40.4 μV); however, an independent-samples *t* test revealed no significant difference [$t_{(67)} = -1.54$, $P = 0.13$, $\eta^2 = 0.03$]. Similarly, the raw startle response magnitude in the neutral condition for the OXT group (25.9 ± 42.7 μV) was not significantly different from that of the PLC group (40.8 ± 34.6 μV) [$t_{(67)} = -1.6$, $P = 0.12$, $\eta^2 = 0.04$].

Experiment 2. ***Behavioral data.*** Two-sample *t* tests revealed no between-group differences in the overall number of subsequently remembered items nor in the numbers of subsequently remembered negative or neutral items (all *P* values > 0.20). Thus, OXT had no influence on memory capacity in general nor on a valence-specific effect. However, we observed a shift in the proportion of subsequently remembered negative to neutral items. Specifically, the OXT group showed a subsequent memory bias toward negative information at the cost of neutral information [remembered negative items minus remembered neutral items: OXT, 4.9 ± 2.8; PLC, −2.8 ± 3.0; $t_{(68)} = -2.98$, $P < 0.01$, $\eta^2 = 0.12$] (Fig. 2*A* and Table S4).

Whole-brain analysis of functional MRI data. The pooled sample showed greater activity in response to negative stimuli in left inferior occipital gyrus [$t_{(34)} = 8.87$, $P_{FWE} < 0.05$, Montreal Neurological Institute (MNI): $x = -44$, $y = -70$, $z = -6$] and right middle occipital gyrus [$t_{(34)} = 8.58$, $P_{FWE} < 0.05$, MNI: $x = 48$, $y = -80$, $z = 4$] (main effect of "valence"). In general, successful encoding was associated with greater responses in left fusiform gyrus [$t_{(34)} = 6.08$, $P_{FWE} < 0.05$, MNI: $x = -30$, $y = -48$, $z = -14$] (main effect of "subsequent memory"). Separate analyses for the PLC and OXT groups yielded no significant clusters of activation for these contrasts. Although there was no main effect of OXT treatment, we found a group × subsequent memory interaction effect in the left anterior insula for negative but not for neutral stimuli [$t_{(68)} = 6.08$, $P_{FWE} < 0.05$, MNI: $x = -38$, $y = 22$, $z = 4$) (Fig. 2*B* and Table S5). Extraction of individual percent signal changes revealed that successful encoding of negative stimuli was associated with greater left anterior insula responses in the OXT group, an effect that was reversed in the PLC group (Fig. 2*B*). Combined masking confirmed that this interaction effect was independent of simple main effects.

Regions of interest analysis of functional MRI data. In line with previous findings (22–25; but see also refs. 26 and 27), amygdala activity was globally reduced as a result of OXT treatment (Table S5). This reduction was histoprobabilistically mapped to the superficial amygdala, reflecting the demonstrated importance of this amygdala subregion in social stimulus processing (34, 35).

Functional connectivity analysis of functional MRI data. Given that the amygdala has been identified as a key target of OXT effects in the

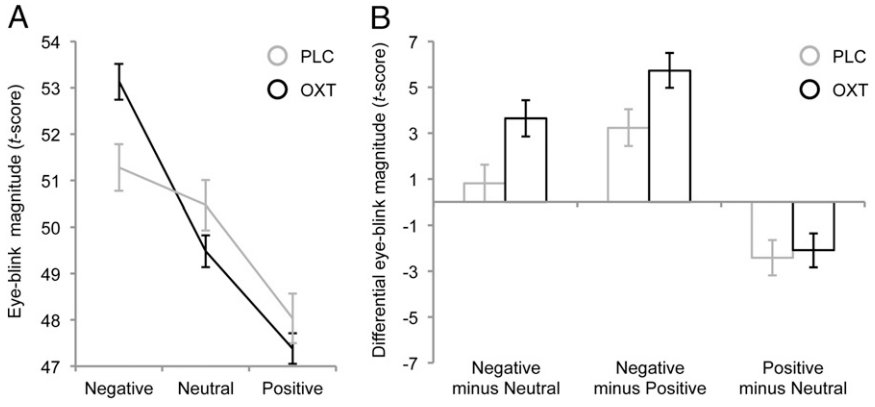

Fig. 1. Exp. 1 examined the effects of OXT relative to PLC on emotion-modulated startle reactivity. (*A*) Both groups showed a significant linear trend for valence (i.e., negative > neutral > positive), but the emotional modulation of eye-blink magnitudes in the OXT group was more than twice as large as in the PLC group. (*B*) The difference in eye-blink magnitude between negative and neutral as well as negative and positive conditions (but not between positive and neutral) was significantly larger in the OXT group than in the PLC group. Taken together, these results show that OXT profoundly potentiated the emotional modulation of the startle response specifically for negative stimuli. Error bars indicate the SEM.

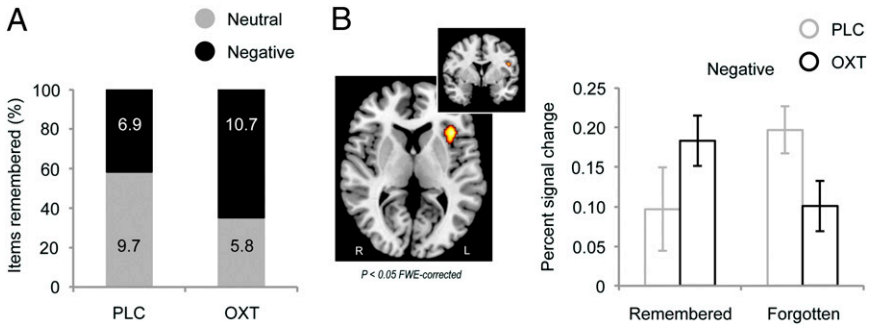

Fig. 2. Exp. 2 examined the effects of OXT relative to PLC on subsequent memory for emotional stimuli. (*A*) Both groups did not differ in their total amount of later remembered items (absolute numbers are given in the bars) in the surprise free-recall test 24 h postscan. However, OXT induced a shift in the relative proportion of later remembered items, evident in a recall bias toward aversive social information. (*B*) Successful encoding of negative stimuli was associated with larger left anterior insula responses (MNI: $x = -38$, $y = 22$, $z = 4$) in the OXT group, an effect that was reversed in the PLC group. Extraction of individual percent signal changes confirmed this pattern. For illustration purposes, results are displayed at uncorrected significance ($P < 0.001$) thresholds. Error bars indicate the SEM.

brain (22–27), we focused, in a first analysis, on potential effects of OXT treatment on functional connectivity of the amygdala using a generalized form of context-dependent psychophysiological interactions (gPPIs) analysis (36). Our results showed that OXT reduced functional coupling between the left amygdala and left anterior cingulate cortex (ACC) [$t_{(68)} = 4.35$, $P_{FWE} < 0.05$, MNI: $x = -4$, $y = 26$, $z = 14$] for all stimuli regardless of valence. For negative stimuli OXT decreased, whereas for neutral stimuli it increased functional coupling between the right amygdala and left insula [$t_{(68)} = 4.23$, $P_{FWE} < 0.05$, MNI: $x = -42$, $y = -2$, $z = 16$]. These profiles were confirmed by extraction of individual parameter estimates (Fig. 3*A*). In a second analysis, we addressed the question of whether the observed OXT effects in the left anterior insula also influenced functional coupling with interconnected brain regions. Our analysis showed that OXT increased functional coupling between left anterior insula and both left inferior frontal gyrus (IFG) [$t_{(68)} = 4.53$, $P_{FWE} < 0.05$, MNI: $x = -46$, $y = 0$, $z = 24$] and left basolateral amygdala (BLA) [$t_{(68)} = 3.46$, $P_{FWE} < 0.05$, MNI: $x = -30$, $y = -14$, $z = -11$] during successful encoding of negative stimuli. Extraction of individual parameter estimates confirmed these profiles in the OXT group, whereas the PLC group exhibited the reverse pattern (Fig. 3*B*). Given a signal distribution of 25% in the hippocampus and 75% in the BLA, histoprobabilistic mapping identified the BLA as the major projection area of OXT-related enhanced cross-talk between the left anterior insula and left amygdala during successful encoding of negative stimuli.

Discussion

Despite the reported anxiolytic effects of OXT, and extensive evidence from our study and others for its suppression of amygdala reactivity to fear signals (22–25; but see also refs. 26 and 27), our results show that the peptide enhances the impact of aversive social information. Specifically, OXT facilitated startle responses to negative stimuli and increased the number of subsequently remembered negative relative to neutral items. Furthermore, our findings suggest that OXT may be acting to promote this potentiated impact of aversive social stimuli by facilitating anterior insula responses to them and by modifying its functional connectivity with the IFG and amygdala. Importantly, despite OXT reducing superficial amygdala responses to negative stimuli, the latter still produced startle and subsequent memory-modulating effects. Histoprobabilistic mapping indeed identified the BLA as the major projection area of OXT-related enhanced cross-talk between the left anterior insula and left amygdala during successful encoding of negative stimuli, suggesting that OXT puts the BLA under neural control of the insula. By hijacking the modulatory functions of the BLA, the insula may become more influential in biasing emotional and cognitive processing toward information gaining crucial behavioral relevance under conditions of enhanced OXT signaling.

Our original aim was to distinguish between three current hypotheses concerning the role of OXT in the human brain, namely as a facilitator of (*i*) prosocial behavior, (*ii*) social salience, and (*iii*) social approach/withdrawal. Our results show that OXT does not simply facilitate responses to positive social cues and reduce them to negative ones, because in both experiments we found evidence for enhanced emotional modulation by, and subsequent memory for, aversive social information. Instead, our data seem to support specific aspects of the social salience and social-approach/withdrawal hypotheses, although before considering which is most relevant, we need to first discuss the implications of our findings in greater detail.

Our result of an enhanced ASR modulation by aversive information following OXT treatment shows that the peptide is not producing an overall anxiolytic effect. Although we found a trend to significant decrease in ASR baseline after OXT application, which is in line with previous findings (22, 37), these anxiolytic actions are clearly not the most influential effects of OXT in our study. Indeed, we found no evidence that OXT affected either arousal or valence ratings for any of the stimuli used. OXT also had no effect on skin conductance (38) or facial electromyographic (EMG) responses during the actual presentation of emotional stimuli (see *SI Text*). Thus, OXT facilitation of ASR modulation by aversive pictures appeared to be independent of any physiological arousal or conscious arousal/valence ratings (39, 40). Although we cannot completely rule out a contribution of physiological arousal changes to the observed facilitating effects of OXT, it seems unlikely that altered conscious feelings of arousal were playing a substantial role.

In view of our findings from Exp. 1, those in Exp. 2 showing an OXT-induced bias toward remembering aversive rather than neutral social information are also unlikely to be a result of altered arousal or valence ratings. Although a number of studies demonstrating anxiolytic effects of OXT (29, 41) and suppression of amygdala reactivity to fear signals (22–25; but see also refs. 26 and 27) have suggested a potential amnestic effect for aversive stimuli, direct evidence for this has been limited. In humans, OXT reduced emotional ratings to aversively conditioned (25) and angry faces (14). Although it has been hypothesized that OXT promotes amnesia for pain, there is in fact no direct evidence for this (1). Thus, despite assumptions that OXT release might aid forgetting of negative emotional experiences, actual evidence in humans is lacking. Indeed, results from our study suggest rather that OXT can bias learning in favor of aversive compared with neutral social information in at least some contexts.

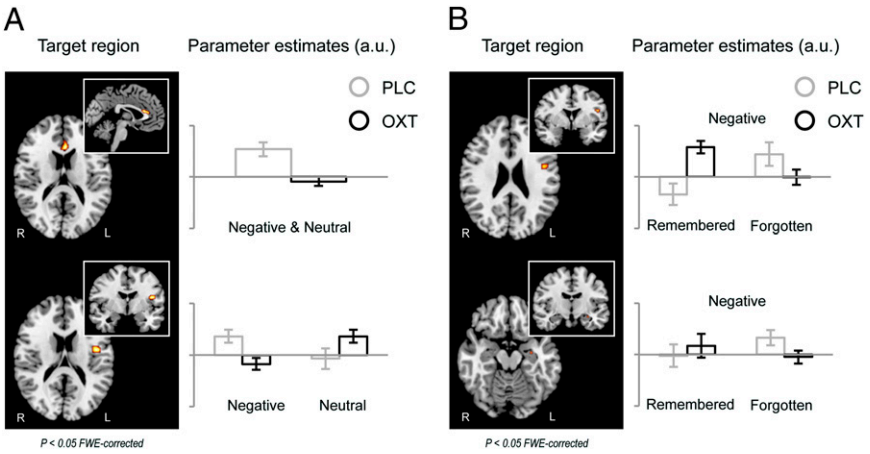

Fig. 3. Results of the functional connectivity analysis of functional MRI data acquired in Exp. 2. (*A*) Effects of OXT on functional connectivity of the left and right amygdala (seed regions). OXT reduced functional coupling between the left amygdala and left ACC (*Upper*) and between the right amygdala and left anterior insula (*Lower*). (*B*) Effects of OXT on functional connectivity of the left anterior insula (seed region). OXT increased functional coupling with the left IFG (*Upper*) and left BLA (*Lower*). Error bars indicate the SEM. For illustration purposes, results are displayed at uncorrected significance ($P < 0.001$) thresholds. a.u., arbitrary units.

Taking these data together, a simple arousal-based explanation of our findings is unlikely. Instead, it seems likely that OXT can enhance some aspect of salience for aversive stimuli independent of arousal. We have shown previously that OXT elevates emotional empathy responses toward pictures of individuals expressing either strong negative or positive emotions (12). In our present study increased activation of the anterior insula for subsequently remembered negative items following OXT treatment may also reflect enhanced empathy and compassion (42), although in this case, empathy and compassion for the pain being suffered by individuals depicted in most of the stimuli. In this context we note that the insula responds both to the actual feeling of pain and also to pain being experienced by someone else (42–44). Although a previous study failed to find any effect of OXT on pain empathy in a paradigm where subjects watched their partner receiving a painful procedure (42), this was arguably a much more salient stimulus than for the subjects in our current study viewing aversive pictures. Indeed, two studies reporting OXT effects on empathy using the "reading the mind in the eyes" (32) or a test of empathic accuracy (45) found that these effects were influenced by task difficulty or initial empathy level. Furthermore, Riem et al. (46) demonstrated OXT-induced decreases of amygdala and increases of insular cortex and IFG responses to crying babies in women without children of their own. The fact that we also found that OXT increased functional coupling between the anterior insula and the IFG provides further evidence of an enhanced empathic response because both regions constitute core areas of the empathy network (42, 47–50). The observed pattern of valence-specific increased anterior insula-IFG coupling thus may reflect enhanced emotional empathic responses (12).

In addition to empathy and compassion, the anterior insula has been implicated in interoceptive awareness and estimation of uncertainty and risk (44, 51–55). In the context of our experiments the effects of OXT in increasing the impact of aversive social information may reflect enhanced perception of visceral reactions and feelings of uncertainty and risk arising from elevated empathic responses, promoting approach and protective behavior toward individuals exposed to social threat.

Our results also reveal an important interplay between the amygdala and the insula after OXT treatment. Not only did OXT reduce superficial amygdala activity but also the functional connectivity between the right amygdala and left insula. Conversely, connectivity between the left insula and left amygdala was strengthened under OXT. Previous research has established amygdala–hippocampal interactions as being key to emotional modulation of memory (56), and the amygdala may also be an important site for OXT modulation of emotional empathy (1, 12). However, our results herein suggest that although OXT suppresses superficial amygdala responses to aversive social stimuli, it may still enhance empathy and modulation of memory in response to these stimuli. This result may be because of the insula taking on a more controlling functional role in mediating these latter behaviors. What would be the advantage of such a dual mechanism of action? One possibility is that it could first promote increased approach behavior toward other individuals exhibiting negative emotional signals but requiring social support (such as crying babies or victims of violence). This result could be achieved by reducing amygdala responses. However, increased caring behavior toward such individuals through enhanced empathy, memory for the emotional event concerned, and heightened preparedness for defense is also required. This result could be achieved by an enhanced insula response. In summary, OXT could alter the balance of interplay between amygdala and insula such that social support and protection is more likely to be offered to individuals who need it, even if they exhibit aversive social signals that would normally be avoided.

Several limitations of our study should be acknowledged. First, our study included only men and, against the background that previous studies reported sex-specific effects of OXT (26, 27), we cannot extrapolate our findings to women. Second, we observed an increased protective/defensive response to social-emotional pictures after OXT administration, but we do not know to what extent these effects are restricted to the social domain.

In summary, OXT clearly does not act simply as an anxiolytic and facilitates recognition of and responses to positive social cues. In terms of the remaining two hypotheses relating to enhancing the salience of social stimuli and increasing approach behavior toward them, our results tend to favor the latter. The absence of altered valence ratings for aversive social stimuli following OXT treatment suggests that increased salience of these stimuli is unlikely because this would also predict altered valence. On the other hand, an increased empathic response toward such stimuli would also predict enhanced approach and potentially protective behavior. However, at the same time our findings of facilitated ASR modulation suggest that in parallel with increased approach behavior, there is also a heightened preparedness for defense or flight. So perhaps a more accurate description of what OXT is promoting in these circumstances is "approach and protective behavior, but with

heightened caution." The empathogenic actions of OXT would clearly facilitate approach behavior but the peptide may also increase preparedness for defense as opposed to flight. The attenuated responsivity of the superficial amygdala to aversive social stimuli may act to reduce the likelihood of flight behavior, whereas the increased insula response may reflect both increased empathy and approach behavior toward threatened and suffering individuals together with a heightened visceral reaction and increased feeling of uncertainty and risk.

Methods

A detailed synopsis of all experimental procedures is provided in the *SI Text*.

Subjects. Detailed information on study participants is provided in the *SI Text*. The present study was approved by the institutional review board of the Medical Faculty of the University of Bonn, registered as a controlled clinical trial (ClinicalTrials.gov Identifiers: NCT01606462 and NCT01607970), and carried out in compliance with the latest revision of the Declaration of Helsinki.

Drug Application. The two experiments underlying the present study followed a randomized, placebo-controlled, double-blind, between-group design; that is, subjects were randomly assigned to either intranasal administration of OXT (24 IU; Syntocinon-Spray, Novartis; three puffs per nostril, each with 4 IU OXT) or PLC (sodium chloride solution), 45 min after which the experimental tasks were carried out.

Exp. 1. Participants were exposed to acoustic startle probes presented either alone or paired with a color picture. The paradigm featured 20 negative, 20 neutral, and 20 positive pictures, presented for 5 s each, which were mostly selected from the International Affective Picture System (57). Similar to the procedure reported by Becker et al. (58), the startle stimulus consisted of a single 50-ms burst of white noise (100 dB) with nearly instantaneous rise and was delivered binaurally via headphones during 60% of the pictures (i.e., 12 from each category) at 2–4 s after picture onset. Facial EMG activity was recorded from two Ag/AgCl electrodes placed over the orbicularis oculi muscle below the left eye (59). Furthermore, electrodermal activity was measured, and facial EMG activity was also analyzed in trials without startle probe. After task completion, participants were administered the self-assessment manikin (57) to obtain behavioral pleasance (valence) and arousal ratings for each picture on a scale ranging from 1 (minimum) to 9 (maximum).

Exp. 2. This experiment consisted of two phases, a functional MRI (fMRI)-scanned encoding phase under drug challenge conditions and a surprise free-recall test 24 h postscan. Critically, for analyzing the subsequent memory (difference due to memory) effect, behavioral responses during free-recall were used to backsort encoding trials as either successful (i.e., subsequently remembered) or not. The fMRI paradigm incorporated 24 aversive and 24 neutral picture stimuli selected from the International Affective Picture System (57) and a verbal description (a noun semantically identical to the picture) (60). Each stimulus was presented four times. Stimuli were shown in a random order for 5 s each, followed by a fixation cross that was presented for 2.5–4.5 s and served as a low-level baseline. MRI data were acquired on a 1.5 Tesla MRI scanner and preprocessed using standard procedures. MRI data were analyzed using SPM8 (Wellcome Trust Centre for Neuroimaging, London, United Kingdom). In the first-level analysis the following conditions were modeled: "neutral: later remembered," "neutral: later forgotten," "aversive: later remembered," and "aversive: later forgotten," using a stick function convolved with a hemodynamic response function (61). To examine if OXT affected the differential neural responses for later-remembered versus forgotten items separately for the aversive and neutral condition, we contrasted the OXT group with the PLC group ($OXT^{\text{aversive: later remembered > aversive: later forgotten}} > PLC^{\text{aversive: later remembered > aversive: later forgotten}}$ and $OXT^{\text{neutral: later remembered > neutral: later forgotten}} > PLC^{\text{neutral: later remembered > neutral: later forgotten}}$). Furthermore, groups were compared using the contrasts "aversive items > neutral items" (main effect of "valence"); "all subsequently remembered items > all forgotten items" (main effect of "subsequent memory"); and "all items > low-level baseline" (main effect of "group"). Emotion-specific effects of OXT treatment were examined using the contrasts "negative items > low-level baseline" and "neutral items > low-level baseline." To address OXT's effects on functional connectivity a PPI analysis was carried out using a gPPI (36). Compared with the standard PPI implementation in SPM, gPPI analysis allows to model more than two task conditions in the same model by spanning the entire experimental space to improve model fit, specificity to true negative findings, and sensitivity to true positive findings (36). First, we examined potential modulatory effects of OXT on functional connectivity of the amygdala. For this aim, we extracted the mean time series for each subject from the left and right amygdala that were defined by using the Wake Forest University Pickatlas (61–63). Second, we examined whether the OXT effect on functional activation in the left anterior insula affected functional coupling of this region with interconnected areas. For this aim, mean time series were extracted from a 6-mm sphere centered at the maximum interaction effect found in the functional activation analysis (MNI: $x = -38$, $y = 22$, $z = 4$). In all analyses, groups were compared using two-sample t tests, and group means were tested using one sample t tests with thresholds of $P < 0.05$ family-wise error-corrected for multiple comparisons. Details on image acquisition, preprocessing and statistical analysis are reported in the *SI Text*.

ACKNOWLEDGMENTS. R.H. was supported by Deutsche Forschungsgemeinschaft Grant HU1302/2-2 and by a Starting Independent Researcher Grant (Neuromodulation of Emotion), jointly provided by the Ministry of Innovation, Science, Research, and Technology of the German State of North Rhine-Westphalia and the University of Bonn.

1. Striepens N, Kendrick KM, Maier W, Hurlemann R (2011) Prosocial effects of oxytocin and clinical evidence for its therapeutic potential. *Front Neuroendocrinol* 32(4): 426–450.
2. Kosfeld M, Heinrichs M, Zak PJ, Fischbacher U, Fehr E (2005) Oxytocin increases trust in humans. *Nature* 435(7042):673–676.
3. Mikolajczak M, et al. (2010) Oxytocin makes people trusting, not gullible. *Psychol Sci* 21(8):1072–1074.
4. De Dreu CK, Greer LL, Van Kleef GA, Shalvi S, Handgraaf MJ (2011) Oxytocin promotes human ethnocentrism. *Proc Natl Acad Sci USA* 108(4):1262–1266.
5. Zak PJ, Stanton AA, Ahmadi S (2007) Oxytocin increases generosity in humans. *PLoS ONE* 2(11):e1128.
6. Fischer-Shofty M, Shamay-Tsoory SG, Harari H, Levkovitz Y (2010) The effect of intranasal administration of oxytocin on fear recognition. *Neuropsychologia* 48(1): 179–184.
7. Marsh AA, Yu HH, Pine DS, Blair RJ (2010) Oxytocin improves specific recognition of positive facial expressions. *Psychopharmacology (Berl)* 209(3):225–232.
8. Schulze L, et al. (2011) Oxytocin increases recognition of masked emotional faces. *Psychoneuroendocrinology* 36(9):1378–1382.
9. Guastella AJ, Mitchell PB, Mathews F (2008) Oxytocin enhances the encoding of positive social memories in humans. *Biol Psychiatry* 64(3):256–258.
10. Savaskan E, Ehrhardt R, Schulz A, Walter M, Schächinger H (2008) Post-learning intranasal oxytocin modulates human memory for facial identity. *Psychoneuroendocrinology* 33(3):368–374.
11. Rimmele U, Hediger K, Heinrichs M, Klaver P (2009) Oxytocin makes a face in memory familiar. *J Neurosci* 29(1):38–42.
12. Hurlemann R, et al. (2010) Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. *J Neurosci* 30(14):4999–5007.
13. Theodoridou A, Rowe AC, Penton-Voak IS, Rogers PJ (2009) Oxytocin and social perception: Oxytocin increases perceived facial trustworthiness and attractiveness. *Horm Behav* 56(1):128–132.
14. Evans S, Shergill SS, Averbeck BB (2010) Oxytocin decreases aversion to angry faces in an associative learning task. *Neuropsychopharmacology* 35(13):2502–2509.
15. Norman GJ, et al. (2011) Selective influences of oxytocin on the evaluative processing of social stimuli. *J Psychopharmacol* 25(10):1313–1319.
16. Shamay-Tsoory SG, et al. (2009) Intranasal administration of oxytocin increases envy and schadenfreude (gloating). *Biol Psychiatry* 66(9):864–870.
17. De Dreu CK, et al. (2010) The neuropeptide oxytocin regulates parochial altruism in intergroup conflict among humans. *Science* 328(5984):1408–1411.
18. Declerck CH, Boone C, Kiyonari T (2010) Oxytocin and cooperation under conditions of uncertainty: The modulating role of incentives and social information. *Horm Behav* 57(3):368–374.
19. Bartz JA, et al. (2010) Effects of oxytocin on recollections of maternal care and closeness. *Proc Natl Acad Sci USA* 107(50):21371–21375.
20. Bartz J, et al. (2011) Oxytocin can hinder trust and cooperation in borderline personality disorder. *Soc Cogn Affect Neurosci* 6(5):556–563.
21. Bartz JA, Zaki J, Bolger N, Ochsner KN (2011) Social effects of oxytocin in humans: Context and person matter. *Trends Cogn Sci* 15(7):301–309.
22. Kirsch P, et al. (2005) Oxytocin modulates neural circuitry for social cognition and fear in humans. *J Neurosci* 25(49):11489–11493.
23. Domes G, et al. (2007) Oxytocin attenuates amygdala responses to emotional faces regardless of valence. *Biol Psychiatry* 62(10):1187–1190.
24. Baumgartner T, Heinrichs M, Vonlanthen A, Fischbacher U, Fehr E (2008) Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron* 58(4): 639–650.
25. Petrovic P, Kalisch R, Singer T, Dolan RJ (2008) Oxytocin attenuates affective evaluations of conditioned faces and amygdala activity. *J Neurosci* 28(26):6607–6615.
26. Domes G, et al. (2010) Effects of intranasal oxytocin on emotional face processing in women. *Psychoneuroendocrinology* 35(1):83–93.
27. Lischke A, et al. (2012) Oxytocin increases amygdala reactivity to threatening scenes in females. *Psychoneuroendocrinology* 37(9):1431–1438.

28. Labuschagne I, et al. (2010) Oxytocin attenuates amygdala reactivity to fear in generalized social anxiety disorder. *Neuropsychopharmacology* 35(12):2403–2413.
29. Heinrichs M, Baumgartner T, Kirschbaum C, Ehlert U (2003) Social support and oxytocin interact to suppress cortisol and subjective responses to psychosocial stress. *Biol Psychiatry* 54(12):1389–1398.
30. Alvares GA, Chen NT, Balleine BW, Hickie IB, Guastella AJ (2012) Oxytocin selectively moderates negative cognitive appraisals in high trait anxious males. *Psychoneuroendocrinology*, 10.1016/j.psyneuen.2012.04.018.
31. Guastella AJ, Mitchell PB, Dadds MR (2008) Oxytocin increases gaze to the eye region of human faces. *Biol Psychiatry* 63(1):3–5.
32. Domes G, Heinrichs M, Michel A, Berger C, Herpertz SC (2007) Oxytocin improves "mind-reading" in humans. *Biol Psychiatry* 61(6):731–733.
33. Kemp AH, Guastella AJ (2012) The role of oxytocin in human affect: A novel hypothesis. *Curr Dir Psychol Sci* 20(4):222–231.
34. Bzdok D, Laird AR, Zilles K, Fox PT, Eickhoff SB (2012) An investigation of the structural, connectional, and functional subspecialization in the human amygdala. *Hum Brain Mapp*, 10.1002/hbm.22138.
35. Goossens L, et al. (2009) Selective processing of social stimuli in the superficial amygdala. *Hum Brain Mapp* 30(10):3332–3338.
36. McLaren DG, Ries ML, Xu G, Johnson SC (2012) A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *Neuroimage* 61(4):1277–1286.
37. Ayers LW, Missig G, Schulkin J, Rosen JB (2011) Oxytocin reduces background anxiety in a fear-potentiated startle paradigm: Peripheral vs central administration. *Neuropsychopharmacology* 36(12):2488–2497.
38. Gamer M, Büchel C (2012) Oxytocin specifically enhances valence-dependent parasympathetic responses. *Psychoneuroendocrinology* 37(1):87–93.
39. Conzelmann A, et al. (2011) Methylphenidate normalizes emotional processing in adult patients with attention-deficit/hyperactivity disorder: Preliminary findings. *Brain Res* 1381:159–166.
40. Thibodeau R (2011) Approach and withdrawal actions modulate the startle reflex independent of affective valence and muscular effort. *Psychophysiology* 48(7): 1011–1014.
41. Quirin M, Kuhl J, Düsing R (2011) Oxytocin buffers cortisol responses to stress in individuals with impaired emotion regulation abilities. *Psychoneuroendocrinology* 36 (6):898–904.
42. Singer T, et al. (2008) Effects of oxytocin and prosocial behavior on brain responses to direct and vicariously experienced pain. *Emotion* 8(6):781–791.
43. Lamm C, Decety J, Singer T (2011) Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage* 54(3):2492–2502.
44. Singer T, Critchley HD, Preuschoff K (2009) A common role of insula in feelings, empathy and uncertainty. *Trends Cogn Sci* 13(8):334–340.
45. Bartz JA, et al. (2010) Oxytocin selectively improves empathic accuracy. *Psychol Sci* 21 (10):1426–1428.
46. Riem MM, et al. (2011) Oxytocin modulates amygdala, insula, and inferior frontal gyrus responses to infant crying: A randomized controlled trial. *Biol Psychiatry* 70(3): 291–297.
47. Amodio DM, Frith CD (2006) Meeting of minds: The medial frontal cortex and social cognition. *Nat Rev Neurosci* 7(4):268–277.
48. Iacoboni M (2009) Imitation, empathy, and mirror neurons. *Annu Rev Psychol* 60: 653–670.
49. Schulte-Rüther M, Markowitsch HJ, Fink GR, Piefke M (2007) Mirror neuron and theory of mind mechanisms involved in face-to-face interactions: A functional magnetic resonance imaging approach to empathy. *J Cogn Neurosci* 19(8):1354–1372.
50. Nummenmaa L, Hirvonen J, Parkkola R, Hietanen JK (2008) Is emotional contagion special? An fMRI study on neural systems for affective and cognitive empathy. *Neuroimage* 43(3):571–580.
51. Rilling JK, et al. (2012) Effects of intranasal oxytocin and vasopressin on cooperative behavior and associated brain activity in men. *Psychoneuroendocrinology* 37(4): 447–461.
52. Ernst J, Northoff G, Böker H, Seifritz E, Grimm S (2012) Interoceptive awareness enhances neural activity during empathy. *Hum Brain Mapp*, 10.1002/hbm.22014.
53. Clark L, et al. (2008) Differential effects of insular and ventromedial prefrontal cortex lesions on risky decision-making. *Brain* 131(Pt 5):1311–1322.
54. Huettel SA (2006) Behavioral, but not reward, risk modulates activation of prefrontal, parietal, and insular cortices. *Cogn Affect Behav Neurosci* 6(2):141–151.
55. Paulus MP, Rogalsky C, Simmons A, Feinstein JS, Stein MB (2003) Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *Neuroimage* 19(4):1439–1448.
56. LaBar KS, Cabeza R (2006) Cognitive neuroscience of emotional memory. *Nat Rev Neurosci* 7(1):54–64.
57. Lang PJ, Bradley MM, Cuthbert BN (2005) *International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual. Technical Report A-6.* (Univ of Florida, Gainesville, FL).
58. Becker B, et al. (2012) Fear processing and social networking in the absence of a functional amygdala. *Biol Psychiatry* 72(1):70–77.
59. Fridlund AJ, Cacioppo JT (1986) Guidelines for human electromyographic research. *Psychophysiology* 23(5):567–589.
60. Hurlemann R, et al. (2005) Noradrenergic modulation of emotion-induced forgetting and remembering. *J Neurosci* 25(27):6343–6349.
61. Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH (2003) An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 19(3):1233–1239.
62. Maldjian JA, Laurienti PJ, Burdette JH (2004) Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *Neuroimage* 21(1):450–455.
63. Tzourio-Mazoyer N, et al. (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15(1):273–289.

Supporting Information

Striepens et al. 10.1073/pnas.1208852109

SI Text

Subjects. ***Sample characteristics.*** Subjects were recruited by local advertisement at the University Bonn and provided written informed consent before study enrollment. Eighty healthy right-handed male volunteers [oxytocin (OXT) group: age range, 19–31, mean = 24 ± 3 y; placebo (PLC) group: age range 21–32, mean = 25 ± 3 y] participated in Exp. 1, and 73 healthy right-handed male volunteers (OXT group: age range 20–25, mean = 26 ± 4 y; PLC group: age range 20–25, mean = 26 ± 4 y) participated in Exp. 2. The subjects were free of current and past physical or psychiatric illness, as assessed by medical history and a Structured Clinical Interview for DSM-IV axis I (SCID-I) and axis II disorders (SCID-II). All participants were naive to prescription-strength psychoactive medication and had not taken any over-the-counter psychoactive medication in the past 4 wk. Participants were asked to maintain their regular bed and wake times and to abstain from caffeine and alcohol intake on the day of the experiment. Tobacco smokers were excluded from participation. In Exp. 2, contraindications for MRI scanning were additional exclusion criteria. In both experiments OXT- and PLC-treated subjects showed no a priori differences regarding age, education, and pretreatment neuropsychological performance (all P values > 0.05) (for details see Tables S1 and S2). To control for potentially confounding effects of OXT on attention and anxiety, all subjects completed the d2 Test of Attention (Aufmerksamkeits- und Belastungstest d2) (1) and the State-Trait Anxiety Inventory (STAI) (2) immediately before the start of the experimental tasks. Analysis of these variables revealed no significant differences between the PLC- and OXT-treated subjects in both experiments (all P values > 0.05) (Tables S1 and S2). Thus, between-group differences cannot be attributed to potential confounding effects of OXT on attention or anxiety. After the experiments, participants were asked to guess whether they had received OXT or PLC; again, there was no significant group difference [Exp. 1: $\chi^2(1) = 0.88$, $P = 0.35$, Cramer-V = 0.12; Exp. 2: $\chi^2(1) = 0.61$, $P = 0.80$, Cramer-V = 0.06].

Neuropsychological screening. To control for pretreatment differences in cognitive performance, all participants completed a comprehensive neuropsychological test battery. Neuropsychological testing included the German version of the RAVLT (Rey Auditory Verbal Learning Test) (3, 4) to assess verbal learning skills, the DST (digit-span test) derived from the revised Wechsler adult intelligence scale (5) to assess working memory performance, the LPS-4 (Leistungspruefsystem Subtest 4) (6) to assess nonverbal reasoning IQ, the MWT-B (Mehrfach-Wortschatz-Intelligenztest Teil B) (7) to assess verbal IQ based on lexical decisions, and the trail-making test (TMT) part A and B (8) to assess visual attention and task-switching abilities. Because a previous study has reported that individual differences in the behavioral inhibition and approach system could account for some of the variance in the emotional modulation of the acoustic startle reflex (ASR) (9), we additionally administered a German version (10) of the questionnaire developed by Carver and White (11) to assess these personality dimensions.

Experiment 1. ***Set-up and stimuli.*** Subjects were seated ∼100 cm in front of a computer screen in a slightly reclined chair with a headrest and instructed to view the pictures presented on-screen and to disregard noises they might hear. The semantic contents of the pictures used in Exp. 1 comprised attractive women (in one picture together with a child) and erotic heterosexual couples in the pleasant condition, household (e.g., knobs, clothespins, a screwdriver) and kitchen objects (e.g., a spoon, a cup) in the neutral condition, and attacking humans (e.g., knife and gun assaults), injured humans (e.g., an accident victim, a starving child, a burning man), and mutilated bodies in the negative condition. Using an in-house programmed script the pictures were adjusted to closely resemble each other in luminance. An independent pilot study involving eight healthy men, none of whom participated in the main experiment, confirmed that the positive and negative pictures were equivalent in terms of arousal ratings. The arousal and valence scores were as follows: negative stimuli (mean = 5.8 ± 0.8 and mean = 2.7 ± 0.6); neutral stimuli (mean = 2.7 ± 1.0 and mean = 5.4 ± 0.4); positive stimuli (mean = 5.0 ± 1.1 and mean = 6.9 ± 0.4). A repeated-measures ANOVA yielded a significant main effect of arousal [$F_{(2, 14)} = 40.16$, $P < 0.01$, $\eta^2 = 0.85$], but Bonferroni-corrected post hoc tests revealed no difference between the negative and positive category ($P = 0.37$). Furthermore, there was a significant linear trend for valence [i.e., negative > neutral > positive; $F_{(1, 7)} = 292.30$, $P < 0.01$, $\eta^2 = 0.98$]. In the ASR paradigm pictures were presented in a pseudorandomized order. Each picture was presented for 5 s. Pictures were presented in two separate runs with 10 pictures of each valence category (30 pictures per run). Pictures were separated by a fixation cross, which was presented in the center of the screen for a randomly generated time interval, ranging from 7 to 17 s (mean, 12 s).

Electrodermal activity. Electrodermal activity (EDA) was measured by two EDA electrodes attached to the thenar and hypothenar of the nondominant hand. The electrodes were filled with a nonhydrating NaCl paste and a constant current of 0.5 V was applied. A commercial system (Contact Precision Instruments) was used for stimulus delivery and psychophysiological recordings. EDA was recorded at a sampling rate of 1,000 Hz. EDA was analyzed for each subject individually for a 10-s period from picture onset. A skin conductance response (SCR) was defined as the first wave in a time window between 1 and 4 s after stimulus onset, with a phasic increase in conductance of more than 0.02 μS. The peak amplitude was baseline-corrected by subtracting the amplitude at stimulus onset. SCR magnitudes were only calculated for trials without startle probe. To normalize the data a log transformation was used. Data from a participant were included in the SCR analysis if at least one nonzero response was available. Eight participants had to be classified as nonresponders and had to be excluded from further analysis. A repeated-measures ANOVA for the SCR magnitude [log(μS + 1)] with "valence" (negative, neutral, or positive) as within-subjects factor and "treatment" (OXT or PLC) as between-subjects factor yielded a significant valence effect [$F_{(2, 118)} = 19.65$, $P < 0.01$, $\eta^2 = 0.25$], but no interaction between valence and treatment [$F_{(2, 118)} = 0.33$, $P = 0.72$, $\eta^2 = 0.01$]. Bonferroni-corrected post hoc tests revealed that both positive (mean = 0.04 ± 0.04, $P < 0.01$) and negative stimuli (mean = 0.03 ± 0.03, $P < 0.01$) elicited a stronger SCR than neutral scenes (mean = 0.01 ± 0.03), with positive pictures even leading to a greater electrodermal response than negative pictures ($P = 0.01$).

Emotion-modulated startle. The startle stimulus consisted of a single 50-ms burst of white noise (100 dB) with nearly instantaneous rise and was delivered binaurally via headphones during 60% of the pictures (i.e., 12 from each category) at 2–4 s after picture onset. A 70-dB white noise background was present throughout the experiment. Facial electromyographic (EMG) activity was recorded from two Ag/AgCl electrodes placed over the orbicularis oculi muscle below the left eye (12). A ground electrode was placed behind the subjects' left ear. A commercial system (Contact Precision Instruments) was used for stimulus delivery and psychophysiological recordings. In addition 18 of 59 interstimulus intervals (ISIs) were

accompanied by startle probes to reduce predictability. To account for early habituation, the experiment started with the presentation of five startle probes in 2-s intervals with no picture and five startle probes during the presentation of a neutral picture. The facial EMG signal was digitized at a rate of 1,000 Hz and amplified with a high-pass filter of 30 Hz and a low-pass filter of 500 Hz. EMG data were rectified and smoothed by a four-point moving average. Startle eyeblink reflex was calculated as the difference between the maximum increase of EMG activity in a time interval between 20 and 100 ms after startle-probe onset and the mean EMG of the 50-ms baseline directly preceding the onset. All EMG data were *z*-transformed within-subjects and then converted into *t*-scores to reduce between-subjects variability and skew. The EMG recordings were visually inspected, and trials with excessive noise were excluded from further analysis. Trials with no perceptible eye-blink reflex were assigned a magnitude of zero and included in the analysis. Subjects displaying fewer than 25% satisfactory blink responses in the paradigm (OXT group, $n = 2$; PLC group, $n = 4$) were excluded. Startle latencies were analyzed in a repeated-measures ANOVA with "group" as between-subjects factor (OXT vs. PLC) and "valence" (negative, neutral, or positive) as within-subjects factor. There was neither a main nor an interaction effect (all *P*s > 0.34). Furthermore, we also analyzed the reaction times for the arousal and valence ratings of the stimuli used in the startle paradigm. The reaction times of all judgments were comparable between both treatment groups, except for the arousal ratings for neutral pictures. In this category, the OXT group was significantly slower than the PLC group [$t_{(59.55)} = 2.09$, $P = 0.04$, $\eta^2 = 0.06$] (Table S3).

Experiment 2. ***Set-up and stimuli.*** We decided to implement a free-recall test for two reasons. First, it has been shown that subsequent memory effects are greater for a free-recall than a recognition test (13). Second, it has been suggested that OXT effects on empathy are more pronounced for difficult compared with easy task items (14), and free-recall tests are invariably more difficult than recognition tests (15). The behavioral data for the free-recall test are presented in Table S4. Each stimulus comprised a colored picture depicting a socially salient situation and its verbal descriptor. Using an in-house programmed script, the pictures were adjusted to closely resemble each other in size and luminance. The verbal descriptor (presented in arial font) was a noun with 4–14 letters that was semantically identical to the picture. Two categories of pictures (negative and neutral) were mainly selected from the International Affective Picture System (16) based on their standard normative scores for emotional arousal and valence. The semantic contents of the pictures used in Exp. 2 comprised daily living situations (e.g., cooking, buying groceries), waiting humans (e.g., on a market, on a street), and working humans (e.g., secretary, musician) in the neutral condition; and attacking or threatening humans (e.g., knife and gun assaults), injured humans or animals (e.g., an accident victim, a starving child, a burning man, extirpation of animals), mutilated bodies, and a snarling dog in the negative condition. In total, 21 of 24 pictures used portrayed human victims of violence/disease. In addition, stimuli were rated by an independent sample of 13 male volunteers on a nine-point self-assessment manikin scale for arousal (1, calm; 9, excited), valence (1, negative; 9, positive), and semantic congruency of the picture and its verbal descriptor (1, lowest; 9, highest). Only stimuli with the highest semantic congruency ($M = 7.1 \pm 1.0$) were included in the final version of the functional MRI (fMRI) paradigm. Arousal and valence scores for the stimuli confirmed the normative emotional arousal and valence ratings (negative stimuli: arousal mean = 6.7 ± 1.2; valence mean = 1.9 ± 0.3; neutral stimuli: arousal mean = 3.1 ± 1.4; valence mean = 6.2 ± 0.6.). Paired-sample *t* tests yielded significant differences between negative and neutral stimuli in arousal [$t_{(12)} = 12.51$, $P < 0.01$] and valence [$t_{(12)} = -23.73$, $P < 0.01$].

fMRI data acquisition and analysis. Stimuli were presented for 5,000 ms and followed by a fixation cross, which served as a low-level baseline. The ISIs ranged between 2,500 ms and 4,500 ms to create jitter. Participants were instructed to fixate the stimuli and to index the valence category of the stimuli (aversive vs. neutral) by button-press. Stimuli were presented to the subjects by means of LCD video goggles (Nordic NeuroLab) connected to a PC running Presentation 14 (Neurobehavioral Systems). fMRI using blood oxygenation level-dependent (BOLD) contrast was carried out on a 1.5 Tesla Siemens Magnetom Espree MRI system (Siemens) using a T2*-weighted echo planar imaging sequence [imaging parameters: TR = 3,000 ms, TE = 50 ms, matrix size: 64 × 64, pixel size: 3 × 3 × 3 mm, slice thickness = 3.0 mm, distance factor = 10%, field of view (FoV) = 210, flip angle = 90°, 35 axial slices]. In addition, high-resolution anatomical images were acquired on the same scanner using a T1-weighted 3D MPRAGE sequence (imaging parameters: TR = 1,660 ms, TE = 3.09, matrix size 256 × 256, pixel size 1×1 mm^2, slice thickness = 1.0 mm, FoV = 256, flip angle = 15°, 160 sagital slices). In total, 950 dynamic scans were recorded, and the task lasted ∼28 min. fMRI data were preprocessed and analyzed using SPM8 software (Wellcome Trust Centre for Neuroimaging, London, United Kingdom; http://www.fil.ion.ucl.ac.uk/spm) implemented in Matlab 7 (MathWorks). The first five volumes of each functional time series were discarded to allow for T1 equilibration. Images were corrected for head movement between scans by an affine registration (17). For realignment, a two-pass procedure was used, by which images were initially realigned to the first image of the time-series and subsequently rerealigned to the mean of all images. For spatial normalization the mean EPI image of each subject was normalized to the current Montreal Neurological Institute (MNI) template (18, 19) using the unified segmentation function in SPM8. This algorithm combines image registration, tissue classification, and bias correction within the same generative model. All images were hereby transformed into standard stereotaxic space and resampled at $2 \times 2 \times 2$ mm^3 voxel size. The normalized images were spatially smoothed using an 8-mm FWHM Gaussian kernel. Raw time series were detrended by the application of a high-pass filter (cutoff period, 128 s). On the first level, the four conditions (neutral: later remembered; neutral: later forgotten; aversive: later remembered; aversive: later forgotten) were modeled by a stick function convolved with a hemodynamic response function (20). The movement parameters were included as confounds in the design matrix. Specific effects were assessed by applying appropriate linear contrasts to the parameter estimates of the experimental conditions resulting in *t*-statistics for each voxel. On the second level, effects of OXT were analyzed comparing the PLC and OXT treated subjects. To examine the modulatory effect of OXT on the difference due to memory (dm) effect for aversive and neutral items, we contrasted the OXT with the PLC group for the contrasts $\mathrm{OXT}^{\text{aversive: later remembered > aversive: later forgotten}} > \mathrm{PLC}^{\text{aversive: later remembered > aversive: later forgotten}}$ and $\mathrm{OXT}^{\text{neutral: later remembered > neutral: later forgotten}} > \mathrm{PLC}^{\text{neutral: later remembered > neutral: later forgotten}}$. In addition, groups were compared for the contrasts "aversive items > neutral items" (main effect of "valence"), "all later remembered items > all forgotten items" (main effect of "subsequent memory"), and "all items > low-level baseline" (main effect of "group"). Emotion-specific effects of OXT treatment were analyzed using the contrasts "aversive items > low-level baseline" and "neutral items > low-level baseline." Group means were tested using one sample *t* tests, and group differences were tested using two-sample *t* tests on the contrasts of interest (OXT > PLC, OXT < PLC) for the whole-brain with a significance threshold of $P < 0.05$ corrected for multiple comparisons based on family-wise error (FWE). The significant interaction effect was again computed and masked with the appropriate main effects thresholded at $P < 0.05$ (uncorrected) to confirm the independence of interaction effects from simple main effects. Given

the pivotal role of the amygdala in emotional memory-encoding and consolidation (21) and the modulatory effects of OXT treatment on amygdala responses to aversive stimuli (22–27), the comparisons were computed repeatedly, this time with an a priori regional focus on the amygdala. In this region-of-interest (ROI) approach the bilateral amygdala was anatomically defined using the Wake Forest University (WFU) Pickatlas (Version 3.0), which provides a method for generating ROI masks based on the Talairach Daemon database (28–30). The implemented atlases are available in MNI space with dimensions of $91 \times 109 \times 91$ sampled at 2-mm intervals, corresponding to the SPM MNI templates. ROI-based two-sample t tests were computed with a threshold of $P < 0.05$ and FWE-corrected for multiple comparisons based on the size of the ROI. The results of these fMRI analyses are presented in Table S5.

Functional connectivity analysis of fMRI data. To address OXT effects on functional connectivity, a psychophysiological interactions (PPIs) analysis was performed. This analysis models condition-dependent changes in connectivity from a chosen seed region to each voxel in the whole brain. We decided to use a generalized form of context-dependent PPIs (gPPIs) (31). Compared with standard PPIs implementation in SPM, gPPIs allows modeling of more than two task conditions in the same PPIs model by spanning the entire experimental space and potentially improves model fit, specificity to true-negative findings, and sensitivity to true-positive findings (31). In a first analysis, we examined the modulation effects of OXT on functional connectivity with the amygdala. For this aim, we extracted the mean time series for each subject from the left and right amygdala defined using the WFU Pickatlas (28–30). Findings from the analysis of functional activation revealed significant effects of OXT treatment in the left anterior insula. To further examine the modulatory effects of OXT on the interplay between this region and other brain regions, mean time series were extracted from 6-mm radius spheres centered at the coordinates of the maximum t-values group × memory interaction effect for negative items in the left anterior insula (MNI: $x = -38, y = 22, z = -4$). Hemodynamic deconvolution was performed on the extracted time series to remove the effects of canonical hemodynamic response function (HRF). The resulting time-series were multiplied by the psychological variables and reconvolved with the HRF to obtain the PPIs interaction terms. gPPIs analysis for each subject was performed on the first level and included regressors for "neutral: later remembered," "neutral: later forgotten," "aversive: later remembered," and "aversive: later forgotten." Separate PPIs models for the three seed regions were computed. On the first level individual contrasts of interest (for the contrasts implemented see fMRI BOLD activation analysis) were computed and submitted to second level two-sample t tests to test for group differences (OXT > PLC; OXT < PLC). Sensitivity to detect OXT effects was increased by restricting between-group comparisons to regions with known functional connections to the amygdala, such as the superior and medial frontal gyri, insula, thalamus, globus pallidus, and anterior cingulate cortex (32). For the anterior insula between-group comparisons were restricted to regions with known functional (33, 34) and structural (35) connections to the anterior insula, namely the superior, middle, and inferior frontal gyri, anterior cingulate cortex, thalamus, and basolateral amygdala. ROIs were defined using the WFU Pickatlas (28–30) and in case of the basolateral amygdala using the Anatomy toolbox (36–38). Groups were compared using two-sample t tests with a significance threshold of $P < 0.05$ and FWE-corrected for multiple comparisons based on the size of the ROI.

Statistics. Demographical, neuropsychological, and psychophysiological data were analyzed using SPSS 19 (SPSS). Quantitative behavioral data were compared by repeated-measures ANOVAs. Partial η^2 was calculated as a measure of effect size. The assumption of sphericity was assessed with Mauchly's test, and for significant violations Greenhouse–Geisser's correction was applied. For qualitative variables, Pearson's χ^2 tests and Fisher's exact tests were used. All reported P values are two-tailed and P values of $P < 0.05$ were considered significant.

1. Brickenkamp R (1995) Aufmerksamkeits-Belastungs-Test 'd2', erweiterte und neu gestaltete Auflage. *Diagnostica* 41:291–296.
2. Spielberger C, Gorsuch R, Lushene R (1970) *Manual for the State-Trait Anxiety Inventory* (Consulting Psychologist Press, Palo Alto, CA).
3. Helmstaedter C, Lendt M, Lux S (2001) *VLMT—Verbaler Lern- und Merkfähigkeitstest [VLMT—Verbal learning and memory test]* (Beltz Test GmbH, Göttingen, Germany).
4. Rey LB (1941) L'examen psychologique dans les cas d'encephalopathie traumatique. *Arch Psychol* 28:55.
5. Wechsler D (1997) *Wechsler Adult Intelligence Scale. Administration and Scoring Manual* (Psychological Corporation, San Antonio, TX), 3 ed.
6. Horn W (1983) *Leistungsprüfsystem L-P-S [Performance testing system L-P-S]* (Hogrefe, Göttingen, Germany).
7. Lehrl S (1978) *Mehrfachwahl-Wortschatz-Intelligenztest MWT-B [Multiple-choice lexical intelligence test MWT-B]* (Dr. Med. Straube, Erlangen, Germany).
8. Raitan RM (1958) Validity of the trail making test as an indication of organic brain damage. *Percept Mot Skills* 8:271–276.
9. Caseras FX, et al. (2006) Influence of individual differences in the Behavioral Inhibition System and stimulus content (fear versus blood-disgust) on affective startle reflex modulation. *Biol Psychol* 72(3):251–256.
10. Hartig JM, Moosbrugger H (2003) The ARES-Scales as a measurement of individual BIS and BAS sensitivity: Development of a long and a short questionnaire version. *Z Differ Diagn Psychol* 24(4):293–310.
11. Carver CS, White TL (1994) Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *J Pers Soc Psychol* 67(2):319–333.
12. Fridlund AJ, Cacioppo JT (1986) Guidelines for human electromyographic research. *Psychophysiology* 23(5):567–589.
13. Paller KA, McCarthy G, Wood CC (1988) ERPs predictive of subsequent recall and recognition performance. *Biol Psychol* 26(1–3):269–276.
14. Domes G, Heinrichs M, Michel A, Berger C, Herpertz SC (2007) Oxytocin improves "mind-reading" in humans. *Biol Psychiatry* 61(6):731–733.
15. Haist F, Shimamura AP, Squire LR (1992) On the relationship between recall and recognition memory. *J Exp Psychol Learn Mem Cogn* 18(4):691–702.
16. Lang PJ, Bradley MM, Cuthbert BN (2005) *International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual. Technical Report A-6* (Univ of Florida, Gainesville, FL).
17. Ashburner J, Friston KJ (2003) *Human Brain Function*, ed Frackowiak RS (Academic, London, UK), 2nd Ed.
18. Evans AC, et al. (1992) Anatomical mapping of functional activation in stereotactic coordinate space. *Neuroimage* 1(1):43–53.
19. Holmes CJ, et al. (1998) Enhancement of MR images using registration for signal averaging. *J Comput Assist Tomogr* 22(2):324–333.
20. Friston KJ, et al. (1995) Statistical parametric maps in functional imaging: A general linear approach. *Hum Brain Mapp* 2(4):189–210.
21. Murty VP, Ritchey M, Adcock RA, LaBar KS (2010) fMRI studies of successful emotional memory encoding: A quantitative meta-analysis. *Neuropsychologia* 48 (12):3459–3469.
22. Kirsch P, et al. (2005) Oxytocin modulates neural circuitry for social cognition and fear in humans. *J Neurosci* 25(49):11489–11493.
23. Petrovic P, Kalisch R, Singer T, Dolan RJ (2008) Oxytocin attenuates affective evaluations of conditioned faces and amygdala activity. *J Neurosci* 28(26):6607–6615.
24. Domes G, et al. (2007) Oxytocin attenuates amygdala responses to emotional faces regardless of valence. *Biol Psychiatry* 62(10):1187–1190.
25. Singer T, et al. (2008) Effects of oxytocin and prosocial behavior on brain responses to direct and vicariously experienced pain. *Emotion* 8(6):781–791.
26. Domes G, et al. (2010) Effects of intranasal oxytocin on emotional face processing in women. *Psychoneuroendocrinology* 35(1):83–93.
27. Gamer M, Zurowski B, Büchel C (2010) Different amygdala subregions mediate valence-related and attentional effects of oxytocin in humans. *Proc Natl Acad Sci USA* 107(20):9400–9405.
28. Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH (2003) An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 19(3):1233–1239.
29. Maldjian JA, Laurienti PJ, Burdette JH (2004) Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *Neuroimage* 21(1):450–455.
30. Tzourio-Mazoyer N, et al. (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15(1):273–289.
31. McLaren DG, Ries ML, Xu G, Johnson SC (2012) A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *Neuroimage* 61(4):1277–1286.
32. Roy AK, et al. (2009) Functional connectivity of the human amygdala using resting state fMRI. *Neuroimage* 45(2):614–626.
33. Cauda F, et al. (2012) Meta-analytic clustering of the insular cortex: Characterizing the meta-analytic connectivity of the insula when involved in active tasks. *Neuroimage* 62 (1):343–355.

34. Cauda F, et al. (2011) Functional connectivity of the insula in the resting brain. *Neuroimage* 55(1):8–23.
35. Flynn FG, Benson DF, Ardila A (1999) Anatomy of the insula—Functional and clinical correlates. *Aphasiology* 13(1):55–78.
36. Eickhoff SB, et al. (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25(4):1325–1335.
37. Eickhoff SB, et al. (2007) Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *Neuroimage* 36(3):511–521.
38. Amunts K, et al. (2005) Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: Intersubject variability and probability maps. *Anat Embryol (Berl)* 210(5-6):343–352.

Table S1. Exp. 1: Demographics and neuropsychological performance

	OXT group	PLC group			
Demographic	Mean (± SD)	Mean (± SD)	*t*	*P*	η^2
Age (y)	24.28 (2.86)	25.03 (2.56)	−1.15	0.26	<0.02
Years of education	16.49 (2.66)	16.85 (1.58)	−0.68	0.49	<0.01
RAVLT					
Trial 1–5	61.72 (5.69)	61.24 (6.45)	0.33	0.74	<0.01
Trial 6 retention	14.19 (1.14)	13.67 (2.30)	1.22	0.23	<0.02
Trial 7 delayed recall	14.50 (0.94)	13.84 (2.29)	1.58	0.12	<0.04
LPS-4	31.28 (3.58)	31.21 (4.55)	0.07	0.95	<0.01
MWT-A	30.08 (3.15)	31.15 (2.95)	−1.45	0.15	<0.03
d2	191.40 (37.11)	189.96 (34.60)	0.01	0.99	<0.01
TMT-A	26.92 (8.24)	24.65 (6.95)	0.95	0.35	<0.01
TMT-B	66.64 (19.32)	60.38 (20.69)	0.91	0.37	<0.02
Digit-span, forward	8.97 (2.00)	8.81 (2.05)	0.36	0.72	<0.01
Digit-span, backward	8.77 (2.27)	9.06 (2.46)	−0.53	0.60	<0.01
BDI	4.81 (4.70)	3.00 (3.48)	1.80	0.08	<0.05
STAI-X1	42.58 (4.09)	41.29 (5.56)	1.09	0.28	<0.02
STAI-X2	43.72 (3.70)	43.91 (4.56)	−0.19	0.85	<0.01

Verbal declarative memory performance was assessed using a German adaption of the RAVLT and included: learning performance across five trials ("Trial 1–5" maximum possible score 75); susceptibility to interference ("Trial 6 retention" maximum possible score 15); and delayed recall ("Trial 7," maximum possible score 15). Nonverbal reasoning IQ was assessed by the LPS-4 (maximum possible score 40). Verbal IQ based on lexical decisions was assessed by the MWT-A (maximum possible score 37); visual attention and concentration was assesses using the d2; visual attention and task-switching was assessed using the TMT-A and TMT-B (results displayed in seconds); working memory performance was assessed using the digit-span forward and backward tests (maximum possible score 14). Depressive symptoms were assessed by the self-report BDI (Beck's Depression Scale, Version II), and anxiety symptoms by the STAI.

Table S2. Exp. 2: Demographics and neuropsychological performance

	OXT group	PLC group			
Demographics	Mean (± SD)	Mean (± SD)	*t*	*P*	η^2
Age (y)	25.53 (4.16)	26.06 (3.54)	−0.57	0.57	<0.01
Years of education	17.59 (3.00)	18.71 (3.38)	−1.47	0.15	<0.03
RAVLT					<0.01
Trial 1–5	59.68 (10.86)	59.29 (7.42)	0.17	0.86	<0.01
Trial 6 retention	12.76 (2.41)	12.51 (2.14)	0.46	0.65	<0.01
Trial 7 delayed Recall	12.79 (2.79)	12.74 (2.14)	0.09	0.93	<0.01
LPS-4	32.56 (3.91)	31.97 (3.83)	0.64	0.53	<0.01
MWT-A	30.56 (2.31)	30.20 (2.44)	0.63	0.53	<0.01
d2	230.09 (37.78)	221.71 (36.11)	0.94	0.35	<0.01
TMT-A	22.97 (6.35)	24.17 (8.89)	−0.65	0.52	<0.01
TMT-B	57.07 (14.68)	61.11 (13.68)	−1.07	0.29	<0.01
Digit-span, forward	8.91 (1.76)	8.63 (1.63)	0.69	0.49	<0.01
Digit-span, backward	8.59 (2.05)	8.20 (1.71)	0.85	0.40	<0.01
BDI	2.68 (3.87)	3.26 (3.34)	−0.67	0.51	<0.01
STAI-X1	41.29 (3.89)	42.97 (4.76)	−1.06	0.29	<0.01
STAI-X2	40.43 (4.49)	42.20 (4.54)	−1.11	0.27	<0.01

Verbal declarative memory performance was assessed using a German adaption of the RAVLT and included: learning performance across five trials ("Trials 1–5," maximum possible score 75), susceptibility to interference ("Trial 6," maximum possible score 15), and delayed recall ("Trial 7," maximum possible score 15). Nonverbal reasoning IQ was assessed by the LPS-4 (maximum possible score 40). Verbal IQ based on lexical decisions was assessed by the MWT-A (maximum possible score 37); visual attention and concentration was assesses using the d2; visual attention and task-switching was assessed using the TMT-A and TMT-B (results displayed in seconds); working memory performance was assessed using the digit-span forward and backward test (maximum possible score 14). Depressive symptoms were assessed by the self-report BDI and anxiety symptoms by the STAI.

Table S3. Exp. 1: Valence and arousal ratings

Valance and arousal	OXT group Mean (± SD)	PLC group Mean (± SD)	*t*	*P*	η^2
Valence					
Negative	2.65 (1.02)	2.66 (0.91)	−0.07	0.95	<0.01
Neutral	4.96 (0.60)	4.85 (0.80)	−0.70	0.49	<0.01
Positive	6.94 (1.37)	7.06 (1.19)	−0.41	0.69	<0.01
Arousal					
Negative	6.19 (1.33)	6.06 (1.39)	−0.41	0.68	<0.01
Neutral	2.04 (0.84)	1.98 (1.08)	−0.26	0.80	<0.01
Positive	5.74 (1.60)	5.82 (1.75)	−0.18	0.86	<0.01
Valence reaction time(s)					
Negative	3.22 (1.09)	3.46 (1.34)	−0.81	0.42	<0.01
Neutral	3.02 (1.05)	2.74 (1.13)	−1.09	0.28	<0.02
Positive	3.16 (1.26)	3.00 (1.07)	−0.57	0.57	<0.01
Arousal reaction time(s)					
Negative	2.90 (0.84)	2.87 (0.85)	−0.15	0.89	<0.01
Neutral	2.46 (0.88)	2.09 (0.55)	−2.09	0.04	<0.06
Positive	2.76 (1.05)	2.80 (1.04)	−0.16	0.88	<0.01

Table S4. Exp. 2: Performance in the surprise free-recall test 24 h postscan

Items	OXT group Mean (± SD)	PLC group Mean (± SD)	*t*	*P*	η^2
Total items remembered	16.46 (7.37)	16.58 (5.96)	−0.07	0.943	<0.01
Negative items remembered	10.66 (3.61)	6.89 (4.01)	−1.29	0.200	0.02
Neutral items remembered	5.8 (2.93)	9.69 (3.94)	−1.08	0.286	0.02
Edm	4.86 (2.78)	−2.80 (2.99)	−2.98	0.004	0.12

Edm, subsequently remembered negative items minus neutral items.

Table S5. Exp. 2: Brain activation differences between oxytocin and placebo groups

Contrast	Hemisphere	MNI coordinates *x*	*y*	*z*	Size*	*P*
Whole-brain analysis						
OXT > PLC						
Later remembered negative > later forgotten negative						
Anterior insular cortex	Left	−38	22	4	14	0.004
Amygdala ROI analysis						
OXT < PLC						
Negative and neutral > baseline						
Superficial subregion	Right	22	−6	−14	8	0.007
Negative > baseline						
Superficial subregion	Right	22	−6	−12	6	0.018
Neutral > baseline						
Superficial subregion	Right	22	−6	−14	7	0.009

*Cluster extent threshold of K (number of contiguous voxels) ≥ 5.