The effects of oxytocin on social reward learning in humans

Rebecca Clark-Elford[1], Pradeep J. Nathan[1,2,3], Bonnie Auyeung[1,6], Valerie Voon[1], Akeem Sule[1,4], Ulrich Müller[1], Robert Dudas[1], Barbara J. Sahakian[1,5], K. Luan Phan[7] and Simon Baron-Cohen[1,6]

[1] *Department of Psychiatry, University of Cambridge, UK*

[2] *New Medicines, UCB Pharma S.A., Belgium*

[3] *School of Psychology and Psychiatry, Monash University, Australia*

[4] *South Essex NHS Partnership Trust, UK*

[5] *MRC/Wellcome Trust Behavioural and Clinical Neuroscience Institute, University of Cambridge, UK*

[6] *Department of Psychiatry, Autism Research Centre, University of Cambridge, UK*

[7] *Departments of Psychiatry and Psychology, University of Illinois at Chicago, USA*

Abstract

It has been hypothesised that the mechanisms modulating social affiliation are regulated by reward circuitry. Oxytocin, previously shown to support affiliative behaviour and the processing of socio-emotional stimuli, is expressed in areas of the brain involved in reward and motivation. However, limited data are available that test if oxytocin is directly involved in reward learning, or whether oxytocin can modulate the effect of emotion on reward learning. In a double-blind, randomised, placebo-controlled, within-group study design, 24 typical male volunteers were administered 24 IU of oxytocin or placebo and subsequently completed an affective reward learning task. Oxytocin selectively reduced performance of learning rewards, but not losses, from happy faces. The mechanism by which oxytocin may be exerting this effect is discussed in terms of whether oxytocin is affecting identity recognition via affecting the salience of happy faces. We conclude that oxytocin detrimentally affects learning rewards from happy faces in certain contexts.

Introduction

The neuropeptide oxytocin is a key modulator of social behaviour. Oxytocin is found across many phyla and has been shown to play a role in social affiliation, specifically pro-social behaviour (Witt et al., 1992), social recognition (Popik et al., 1992), and pair bonding (Insel and Hulihan, 1995). Moreover, oxytocin supports affiliative behaviour in humans by increasing social behaviours such as trust, generosity and empathy (Kosfeld et al., 2005; Domes et al., 2007a; Zak et al., 2007). Oxytocin receptors are found throughout the mammalian brain and, of particular interest, areas of the brain that are associated with motivation and reward, namely the amygdala (Huber et al., 2005), nucleus accumbens (NAcc) and ventral striatum (Buijs et al., 1985; Lee et al., 2009). A key neural structure implicated in reward learning is the mesolimbic system: a configuration of dopaminergic neurons that originates in the ventral tegmental area and terminates in the NAcc, amygdala and hippocampus. Animal models also indicate that these areas of the brain are important in the formation of pair bonds in monogamous mammals (Liu and Wang, 2003). Moreover, administration of oxytocin can induce pair bonding behaviours whilst oxytocin antagonists can block such mating preference (Insel and Hulihan, 1995). Complimentary evidence also suggests that both oxytocin and dopamine are necessary for such partner formation and that these interactions are regulated in the NAcc (Young et al., 2001; Liu and Wang, 2003). Therefore, it could be hypothesised that the mechanisms modulating affiliation and social interaction may be regulated by neural reward circuitry (Young et al., 2001).

Despite evidence demonstrating significant overlap in oxytocin receptor distribution and areas of the brain involved in reward circuitry, limited evidence suggests that oxytocin is directly involved in social reward learning in humans. Moreover, whether oxytocin can affect how emotion interacts with learning rewards or punishment from social stimuli. Oxytocin is heavily involved in the processing of negative social stimuli in humans (Kirsch et al., 2005; Petrovic et al., 2008; Labuschagne et al., 2010) and increasing evidence demonstrates

Address for correspondence: R. Clark-Elford or Pradeep J. Nathan, Department of Psychiatry, University of Cambridge, UK.
Tel.: +44 (0) 1223 336583   *Fax*: +44 (0) 1223 336968
Email: rjc84@cam.ac.uk or pn254@cam.ac.uk

the pro-social effects of oxytocin and processing of happy expressions. For instance, oxytocin can increase recognition of happy facial expressions (Marsh et al., 2010; Schulze et al., 2011), increase processing of positive social cues (Di Simplicio et al., 2009), increase the response of the amygdala to happy faces (Gamer et al., 2010) and potentiate the effects of social reinforcement learning (Hurlemann et al., 2010).

One study concluded that although oxytocin is not directly involved in financial reward learning, on administering a social associative learning task, oxytocin did reduce aversion to angry faces when previous financial feedback strongly favoured to do so (Evans et al., 2010). This study, although demonstrating that oxytocin interacts with social reward learning, was constrained by a number of factors. The task design was such that two emotive faces were paired (an angry and happy, or sad and happy) and participants were instructed to optimise financial gains by selecting the face that was more likely to lead to a financial reward. In pairing two different highly emotive faces with no neutral condition it is difficult to clarify the interaction between reward learning and emotional processing and how this interaction is modulated by oxytocin. Although oxytocin reduced aversion to angry faces, it could have equally reduced the salience of the appetitive attributes of happy faces. By including a neutral condition, the interaction of emotion and reward learning can be investigated and whether this is modulated by oxytocin. Additionally, this study did not include a specific 'punishment' condition. By including a loss condition, interactions of emotion can be identified separately for reward and punishment learning, and can also identify whether oxytocin has specific effects on either type of learning.

Therefore, the primary aims of this study were to assess how emotion interacts with social reward and punishment learning, and whether oxytocin modulates these effects in typical volunteers. Oxytocin has been shown to be involved in processing of positive social stimuli (Di Simplicio et al., 2009; Gamer et al., 2010; Marsh et al., 2010; Schulze et al., 2011) and given the overlap in receptor distribution with reward learning circuitry (Buijs et al., 1985; Young et al., 2001), we hypothesised that oxytocin would interact with emotion to affect processing speed and learning from happy faces in rewarding situations. Specifically, we predicted that oxytocin would reduce reaction times to happy faces, and in doing so affect performance in learning rewards from happy faces. Given that participants may find more attractive faces rewarding and that oxytocin has been shown to increase ratings of facial attractiveness (Theodoridou et al., 2009) an 'attractiveness questionnaire' was utilised. Specifically, we examined whether oxytocin modulates ratings of facial attractiveness during reward learning. We hypothesised that oxytocin will interact with emotion to affect ratings of facial attractiveness of the actors used in the task.

Method

Participants

A total of 27 males, average age 26.28 yr (S.D.=6.29, age range 18–42 yr), were recruited for this study via the Internet and advertisements in the University of Cambridge and surrounding area. Participants were excluded from participation if they had any current or previous history of DSM-IV Axis I disorders [as verified by the Mini International Neuropsychiatric Interview (Sheehan et al., 1998), administered by a trained psychiatrist], any other physical medical condition, history of alcohol or substance abuse (within 12 mth of study entry), smoked and/or were currently prescribed medication known to affect brain function. Participants provided written informed consent and the study was conducted in accordance with the UK National Research Ethics Committee (REC Reference: 10/H0308/77) and NHS Research and Development guidelines.

Study design

This study utilised a double-blind, randomised, placebo-controlled, within-group (crossover) study design whereby participants were tested under two acute treatment conditions separated by at least a 1 wk wash-out period. Treatment conditions included an active intranasal oxytocin spray (24 IU, 40.32 μg, Syntocinon-spray; Novartis, Switzerland), three actuations were administered to each nostril, (4 IU, 6.72 μg each), and a placebo (containing all ingredients except for the peptide).

Procedure

Each participant attended three separate sessions, a screening session followed by two nasal spray study visits. During the screening session, participants were assessed using the Mini International Neuropsychiatric Interview (Sheehan et al., 1998), administered by a trained psychiatrist. This interview was administered to ensure that participants did not suffer from any DSM-IV Axis I disorders. Participants were also administered the Beck Depression Inventory- II (Beck et al., 1996) and the National Adult Reading Test (Monk et al., 2006) to assess current levels of mood and IQ.

Prior to study visits all participants were asked to refrain from caffeine on the day of testing and alcohol was not permitted 24 h prior to the session. Participants completed two pre-drug questionnaires: the Bond and Lader Visual Analogue Mood Scale [VAS (Bond and Lader, 1974)] and the Spielberg State-Trait Anxiety Index [STAI (Spielberg et al., 1983)], to assess mood and anxiety. Following a brief medical assessment, nasal sprays were self-administered, supervised by a medical professional and participants were then advised to rest for 45 min, consistent with previous studies (Kirsch et al., 2005; Kosfeld et al., 2005; Domes et al., 2007a; Zak et al., 2007; Guastella et al., 2008b; Andari et al.,
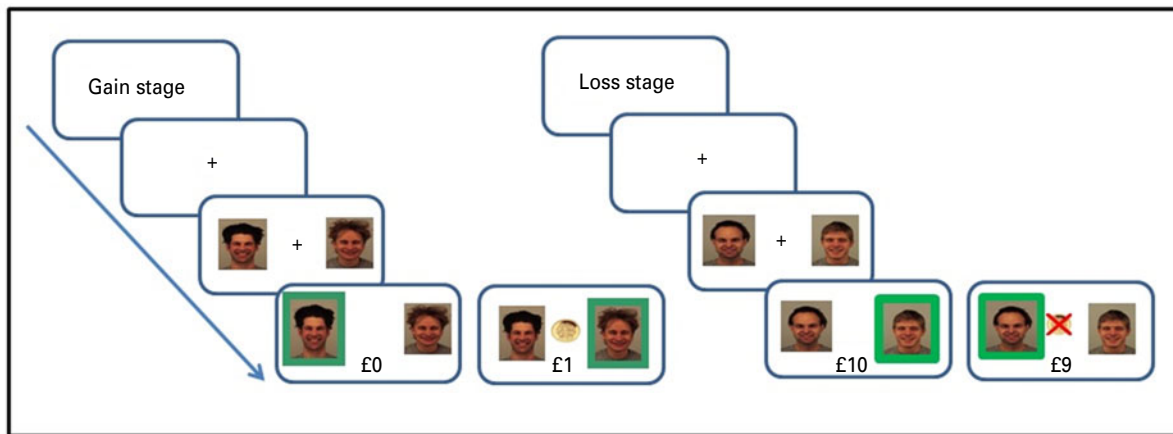
Fig. 1. Illustrative outline of ART.

2010), and physiological evidence (Born et al., 2002). After this break, participants were again medically assessed and task administration began. This study formed part of a larger study that involved three behavioural tasks. Following cognitive testing post-drug questionnaires (VAS and STAI) were administered as well as an 'attractiveness questionnaire'. Participants were also asked whether they thought they received oxytocin or placebo during the session.

Experimental task

Social reward and punishment learning were assessed using the Affective Reward Task (ART; similar to Evans et al., 2010) designed specifically to assess social reward learning from financial feedback. This task involved presenting face-pairs to participants during a 'gain' block and a separate 'loss' block. Face-pairs were constructed using a validated set of face stimuli (Lundqvist et al., 1998), and consisted of two faces expressing the same emotion where the actors had different identities. Each face within the pair was assigned a contingency (80:20), such that during the 'gain' block one face on 80% of trials would lead to a financial reward (£1) whilst the other face would lead to a financial reward on 20% of trials. If participants received a reward a £ sign would appear in the centre of the screen and £1 would be added to their pot. If participants failed to receive a reward, nothing was displayed on screen. During the 'loss' block one face would lead to a loss on 80% of trials, whilst the face other would lead to a loss on 20% trials. If participants lost, a £ with a cross through it was displayed and £1 was lost from their pot, otherwise nothing was displayed. Participants were instructed to optimise their winnings by gaining as much money as possible in the 'gain' block. In the 'loss' block participants were given £50 and instructed to avoid losing money. Participant indicated their choice via a button press. Participants were presented with happy, fearful and neutral face-pairs, 26 presentations of each face-pair. Participants were instructed that they

would win a proportion of their winnings, to a maximum of £5. However, in reality all participants were compensated the full £5. See Fig. 1 for an illustrative outline of ART.

Drug administration for the participants included in the final analysis was fully counterbalanced, with half receiving oxytocin at their first session. ART was also counterbalanced for gender and order of blocks. To counterbalance for order of blocks; half the participants, who received oxytocin first, completed the 'gain' block followed by 'loss', and vice versa. To counterbalance for gender, half the participants, who received oxytocin first, were shown male faces, the remaining half were shown female faces. The counterbalancing was conducted in the same way for those who received the placebo at their first visit.

The order of block and gender remained constant for both sessions. However, participants were shown different actors at each session. The ART programme randomly allocated a picture to blocks (either 'gain' or 'loss'), and then to contingency (80 or 20%) from the available stimuli in the chosen photo set. Face-pair run order was also randomly assigned, however, no face-pair could be presented more than three times consecutively. Faces were also presented equally on the left and right hand side of the screen. As this task formed part of a larger cognitive battery, the order of tasks was also counterbalanced across participants using a Latin-squared design, but remained constant across sessions.

ART (V1.2.2.0) was programmed using Visual Basic and was presented on a Sahara Slate PC i400 Series. A Cedrus RB-830 Response Pad was utilised for participants to indicate left or right. Following cognitive testing an 'attractiveness questionnaire' was administered to all participants. Participants were instructed to indicate how attractive they thought each of the actors in the ART task were from 1 (very attractive) to 9 (very unattractive). Participants rated all 16 actors used during their session and all actors were depicted with a neutral expression.

Statistical analysis

The main dependent variables for ART include several performance measures including 'percentage correct' across all trials within a condition, 'peak learning', which includes percentage correct for trials 2–12 and 'trials to criterion', which is indicative of how many trials it took for participants to get four consecutive correct responses. For the purpose of this study 'correct' is defined as choosing the face assigned to an 80% contingency in the 'gain' block and 20% contingency in the 'loss' block: i.e. they were choosing the face that usually led to optimal financial outcomes. Reaction times for correct scores within criterion were also assessed as a dependent variable; here a criterion was applied such that responses were excluded if a participant's reaction time on individual trials was 2 S.D. above/below his own mean.

For all dependent variables, if the data were normally distributed a 2 (oxytocin, placebo)×2 (gain, loss)×3 (happy, fearful, neutral) repeated-measures ANOVA was used. If main interactions were significant, post hoc pairwise comparisons were made and a Bonferroni correction was implemented for multiple comparisons. If data were non-normally distributed, a Wilcoxon Signed Rank Test or a Friedman's two-way analysis of variance was utilised, as indicated, based on results of the main ANOVA's for reaction time and overall 'percentage correct' data. Scores on the 'attractiveness questionnaire' were all normally distributed and a 2 Drug (oxytocin, placebo)×2 Monetary (gain, loss)×3 Emotion (happy, fearful, neutral)×2 Contingency (80, 20%) repeated-measures ANOVA was used. Again, if main interactions were significant, post-hoc pairwise comparisons were made and a Bonferroni correction was implemented for multiple comparisons.

Changes in subjective mood and anxiety (VAS and STAI) were assessed twice, pre- and post-drug administration. The VAS (Bond and Lader, 1974) consists of 16 visual analogue scales and participants marked on a 10 mm line according to how they felt at the moment. Items were combined into 3 main factors: alertness, contentedness and calmness. Separate analyses were conducted for each factor and changes in scores over session and time were assessed using a 2 Drug (oxytocin, placebo)×2 Time (pre-, post-) repeated-measures ANOVA. The STAI (Spielberg et al., 1983) consists of two parts. The 'state' measures how anxious the participant is feeling at that moment and participants rate on a four-point likert scale, (1 being not at all, 4 being very much so). The 'trait' measures how anxious participants are feeling in general, again participants rate on a four-point likert scale. Separate analyses were conducted for 'state' and 'trait' anxiety and changes in scores over session and time were assessed using a 2 Drug (oxytocin, placebo)×Time 2 (pre-, post-) repeated-measures ANOVA.

Results

Demographics and mood and anxiety questionnaires

Demographics and data from the mood and anxiety questionnaires are reported in Supplementary Table S1. Data from 24 participants was included in the final analysis as one participant withdrew from participation, data was not successfully recorded for one participant and one participant was excluded as an outlier who demonstrated continuous anticipatory responses and failed to learn contingencies. There were no effects of oxytocin on the VAS or STAI. Repeated-measures ANOVAs for the VAS did not reveal any main effects of drug for alertness ($F(1,23)=1.95$, $p=0.18$), calmness ($F(1,23)=1.37$, $p=0.25$) or contentedness ($F(1,23)=2.19$, $p=0.15$). There were also no drug×time interactions for alertness ($F(2,46)=0.12$ $p=0.73$), calmness ($F(2,46)=0.019$, $p=0.89$), or contentedness ($F(2,46)=0.261$, $p=0.61$). Repeated-measures ANOVAs for the STAI revealed no main effects of drug on state ($F(1,23)=0.90$, $p=0.35$) or trait ($F(1,22)=3.88$, $p=0.06$) anxiety and there were also no significant drug×time interactions for state ($F(2,46)=0.301$, $p=0.59$) or trait ($F(2,44)=1.17$, $p=0.29$) anxiety. A t-test also revealed that participants were no better at guessing which spray they received at either the oxytocin or placebo session, with 62.5% of participants guessing correctly at each session ($p=1.00$).

Reaction time data

Reaction time data for correct responses, within criterion, are depicted in Fig. 2. The repeated-measures ANOVA revealed a main effect of the monetary condition, such that overall participants were faster in the 'gain' condition than the 'loss' condition, ($F(1,23)=13.112$, $p=0.001$). Analysis also revealed a main effect of emotion ($F(2, 46)=13.510$, $p<0.000$), such that participants were significantly faster for neutral faces than happy faces ($p=0.024$) and participants were also significantly faster for neutral faces than fearful faces ($p<0.000$). Further analysis revealed a monetary×emotion interaction, $F(2,46)=3.477$, $p=0.039$, such that participants were significantly slower in fearful trials during the 'loss' condition compared to the 'gain' condition ($p<0.0001$). However, there were no significant differences for happy ($p=0.277$) and neutral faces ($p=0.103$) comparing 'gain' and 'loss' conditions. Finally, as demonstrated in Fig. 2d, it would seem that participants were particularly slower for happy faces post-oxytocin, than post-placebo ($p=0.022$). However, the main drug×emotion interaction was not significant, $F(2, 46)=2.213$, $p=0.121$.

Performance data: overall percentage correct

Overall percentage-correct data is depicted in Fig. 3. This data was non-normally distributed and was subsequently Log10 transformed. This transformation successful normalised the data and a repeated-measures ANOVA
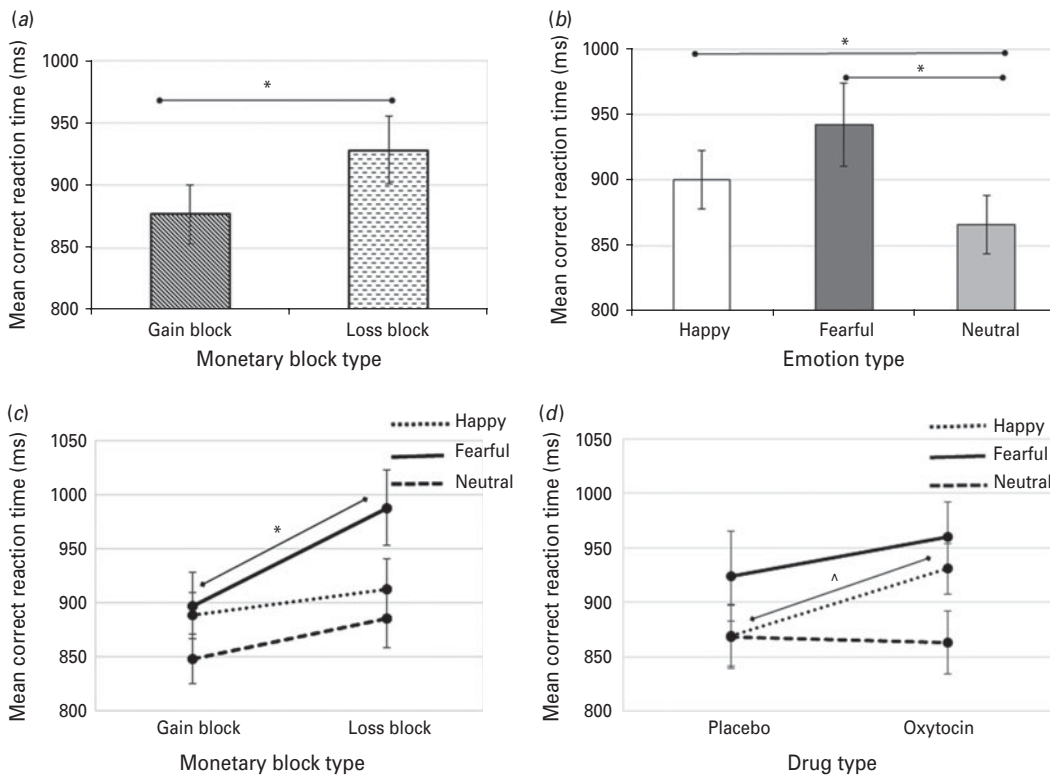
Fig. 2. Mean reaction time data, within criterion: a. Difference in mean reaction time between the 'gain' block and 'loss' block, b. Differences in mean reaction between happy, fearful and neutral trials. c. Depicts the interaction of monetary and emotion and d. Depicts the potential interaction of drug × emotion. * *indicates a significant difference (p<0.05) and ^ indicates a significant pairwise comparison (p<0.05). However, the main interaction is ns.*

revealed no main effects of drug, emotion or money. However, analysis did show a significant drug × emotion interaction, $F(2,46)=4.238$, $p=0.020$. *Post-hoc* comparisons revealed that post-oxytocin participants had significantly worse performance for happy trials than neutral ($p=0.005$). This effect was not apparent post-placebo ($p=1.000$).

Performance data: trials to criterion

Trials-to-criterion data are depicted in Fig. 4. This data was non-normally distributed and remained so after transformation. Therefore, to assess learning rewards from happy faces, Wilcoxon Signed Rank Tests revealed that participants did demonstrate significant decreased learning from happy faces, in the 'gain' block, for oxytocin compared to placebo, $T=232.500$, $p=0.001$. Participants, however, demonstrate no significant differences in learning from fearful faces, in the 'gain' block, for oxytocin compared to placebo ($p=0.614$, Fig. 4a). Furthermore, participants demonstrated no significant differences in learning from happy ($p=0.970$) or fearful faces ($p=0.651$) in the 'loss' block, for oxytocin compared to placebo (Fig. 4b). To examine learning from happy faces averaging across 'loss' and 'gain' blocks, a Friedman's two-way ANOVA revealed no significant differences in learning after oxytocin administration,

compared to placebo, across all emotive conditions ($p=0.073$, Fig. 4c).

Performance data: peak learning

Peak-learning data is depicted in Fig. 5. The data was non-normally distributed and remained so after transformation. Therefore, to assess learning rewards from happy faces Wilcoxon Signed Rank Tests revealed that participants did demonstrate significant decreased learning from 'gain' faces, in the reward block, for oxytocin compared to placebo, $T=1.000$, $p<0.000$. However, participants demonstrated no differences in peak learning from fearful faces in the 'gain' block, for oxytocin compared to placebo ($p=0.808$; Fig. 5a). Furthermore, there were no significant drug differences found for learning from fearful ($p=0.758$) or happy ($p=0.986$) faces in the 'loss' condition (Fig. 5b).

To examine learning from happy faces averaging across 'loss' and 'gain' blocks, a Friedman's two-way ANOVA revealed significant differences in learning after oxytocin administration, compared to placebo, $X^2=26.674$, $p<0.000$. Pairwise comparisons, using a Bonferroni corrected alpha level of 0.0167, demonstrate that participants show decreased learning from happy faces post-oxytocin than placebo, compared to neutral ($T=1.333$, $p<0.000$) and also decreased learning from happy faces
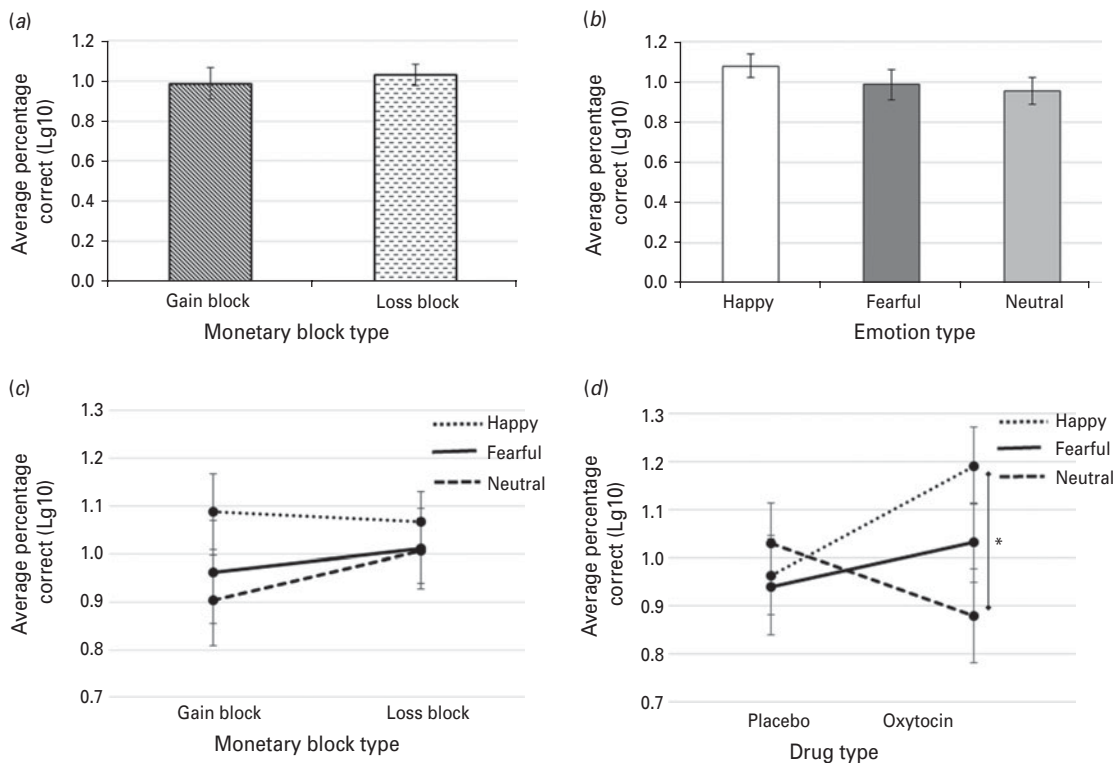
Fig. 3. Overall percentage correct data: a. Difference in 'overall percentage correct' between the 'gain' block and 'loss' block, b. Differences in performance between happy, fearful and neutral trials. c. Depicts the interaction of monetary and emotion and d. Depicts interaction of drug×emotion. * *indicates significant difference (p<0.05). Note reversed Log10 transformed. Therefore, higher scores indicate worse performance.*

post-oxytocin than placebo compared to fearful ($T=1.229$, $p<0.000$, Fig. 5c).

'Attractiveness questionnaire' data

Missing data are noted for this questionnaire. One participant was excluded from the ART data for outlier results and a further three participants were excluded due to missing data. Therefore a total of 22 participants were included in the final analysis (see Supplementary Figure S2). A repeated-measure ANOVA revealed a main effect of contingency, $F(1,21)=6.192$, $p=0.021$: participants rated those actors who were previously presented as given a 20% contingency as significantly less attractive than those given an 80% contingency. Analysis also revealed an emotion×contingency interaction, $F(2,42)=3.621$, $p=0.035$: for those actors previously assigned a 20% contingency, participants rated those actors presented as happy as significantly more attractive than those actors previously presented as neutral ($p=0.036$), using a Bonferroni correction. Analysis also revealed a significant drug×emotion×contingency interaction, $F(2,42)=4.244$, $p=0.021$: for happy faces that were assigned an 80% contingency, participants rated these actors as significantly less attractive post-oxytocin, than post-placebo ($p=0.039$), using a Bonferroni correction.

Discussion

There is significant overlap in oxytocin receptor distribution and areas of the brain involved in reward circuitry, but there is limited evidence that oxytocin is directly involved in reward learning in humans. This study addressed whether oxytocin modulates the effects of emotion on reward or punishment learning in typical volunteers. Results indicate that post-placebo administration there are no differences in learning from happy faces compared to neutral faces. However, post-oxytocin, participants demonstrate reduced learning for happy faces in both the 'gain' and 'loss' block. Furthermore, considering trials of peak learning and the number of trials it takes participants to learn the contingencies, post-oxytocin participants demonstrate reduced learning for happy faces in the 'gain' but not the 'loss' block. To our knowledge, this is the first study to suggest that oxytocin negatively impacts learning from happy faces, especially when associated with rewarding conditions. We discuss potential mechanisms that could explain these effects.

Reduced learning from happy faces in rewarding conditions following oxytocin may seem surprising, given the evidence demonstrating the pro-social effects of oxytocin. Therefore, it is interesting to consider the mechanisms by which oxytocin may be exerting effects in this
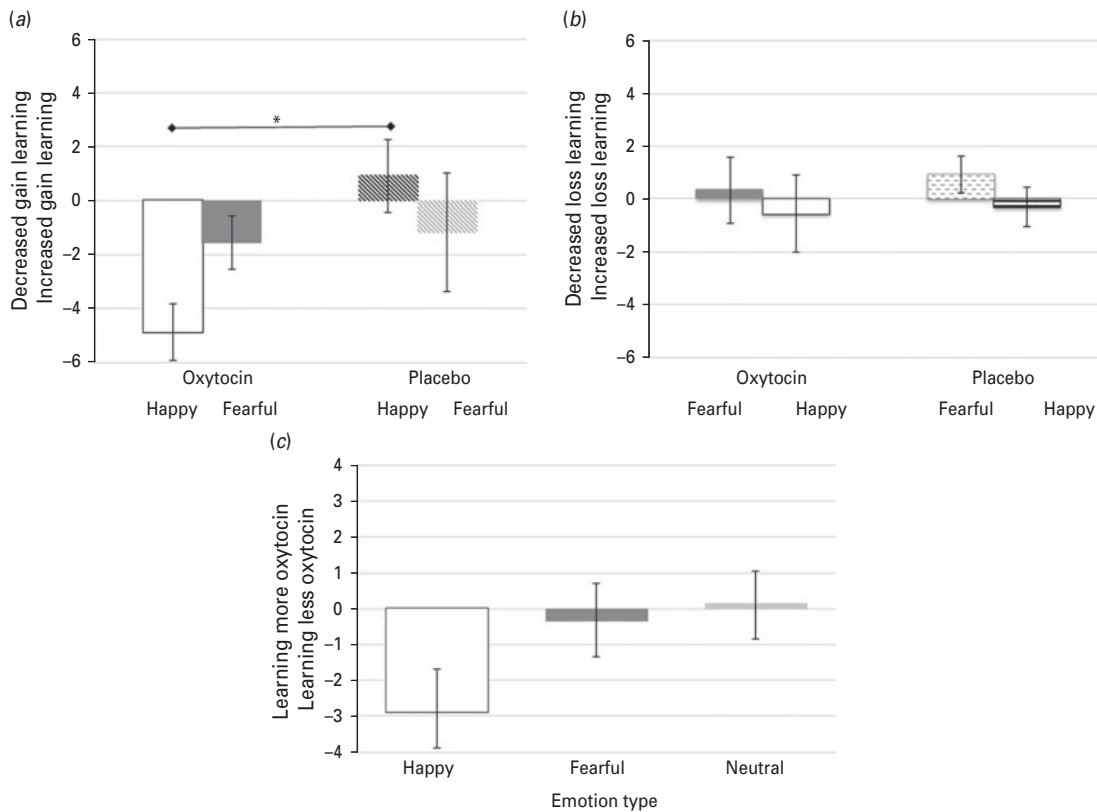
Fig. 4. Trials to criterion data: a. Comparison of oxytocin *vs.* placebo during the gain condition for happy trials and fearful trials (neutral-happy, neutral-fearful), b. Comparison of oxytocin *vs.* placebo during the loss condition for fearful and happy trials (neutral-fearful, neutral-happy), c. Comparison of mean differences in drug effect for each emotion condition (collapsed across gain and loss conditions). *indicate significant differences.*

task. As opposed to Evans et al. (2010), both faces in a pair expressed the same emotion. Therefore, it is assumed that participants employed identity recognition to successfully complete this task. Considering evidence that oxytocin can increase recognition of happy expressions (Schulze et al., 2011), increase processing of positive social cues (Di Simplicio et al., 2009), increase gaze to the eye region (Guastella et al., 2008a) and increase the response of the amygdala to happy faces (Gamer et al., 2010), it could be hypothesised that oxytocin selectively increases the salience of happy faces in rewarding conditions. Furthermore, evidence suggests that emotion cues are processed prior to facial identity, see review (Calder and Young, 2005), therefore oxytocin may bias processing of emotional cues, which could have subsequently impacted negatively on processing of identity. (Guastella et al., 2008a)

Indeed, given that facial emotion can affect the processing of identity (D'Argembeau and Van der Linden, 2011), it is not surprising that if oxytocin increases the salience of happy faces, this is likely to impact negatively on the processing of identity. This is substantiated by our results demonstrating that post-oxytocin, learning from neutral faces resulted in the best performance. Furthermore, our results are also consistent with evidence suggesting that oxytocin increases identity recognition for

angry and neutral faces, but not happy faces (Savaskan et al., 2008).

It is beyond the scope of this study to test whether oxytocin directly effects the processing of happy faces in rewarding conditions. This is a limitation to the study and could have been tested by examining whether participants were impaired in discriminating between happy faces in the absence of a reward component. However, our evidence is substantiated by results demonstrating that oxytocin selectively detriments learning from happy faces in rewarding, but not punishing conditions. If the effects of oxytocin were primarily due to an enhanced bias toward happy, faces learning would been impaired in both reward and punishment conditions. Although ventral striatal circuits have been widely accepted as involved in the processing of reward, whether these circuits are also implicated in punishment, or whether punishment learning relies on a distinct system is still debated. Evidence suggests that punishment learning involves areas of the brain including the dorsal striatum and insula (Mattfeld et al., 2011; Palminteri et al., 2012). It is also noteworthy that although the insula has low expression levels of oxytocin receptors, the dorsal striatum does not show a similar expression of oxytocin receptors as do the ventral striatum (Gimpl and Fahrenholz, 2001). If oxytocin receptors are not expressed in areas
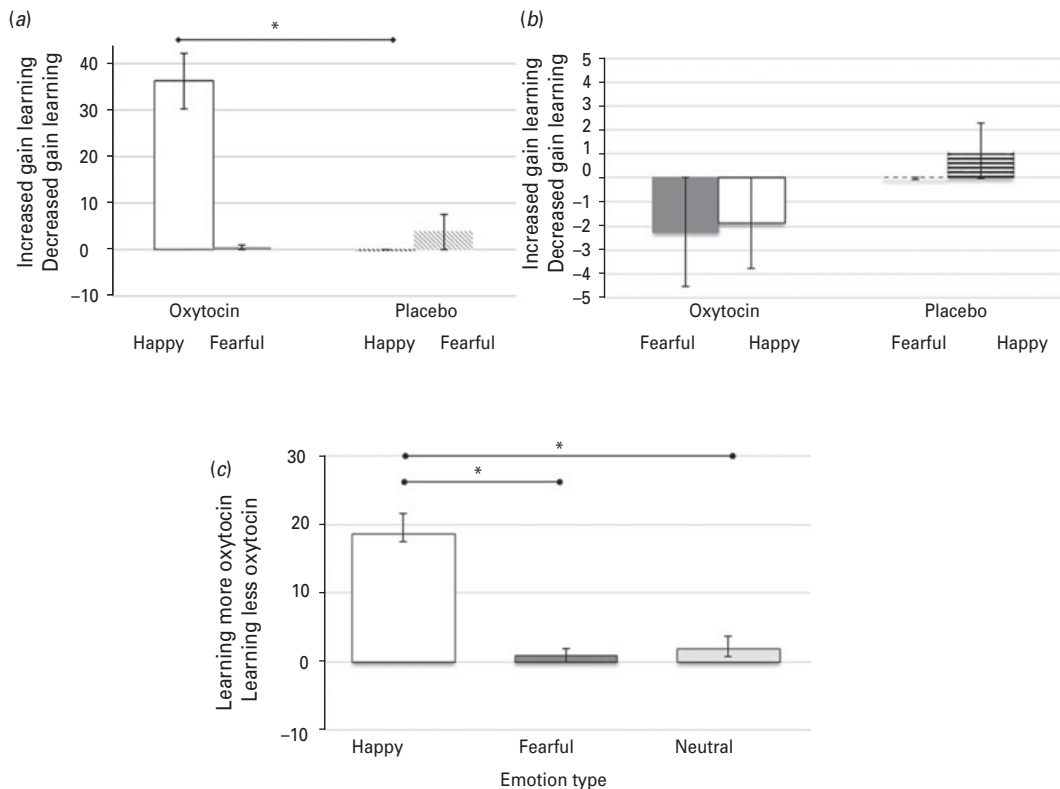
Fig. 5. Peak learning (Average percentage correct for trials 2–12): a. Comparison of oxytocin *vs.* placebo during the gain condition for happy trials and fearful trials (neutral-happy, neutral-fearful), b. Comparison of oxytocin *vs.* placebo during the loss condition for fearful and happy trials (neutral-fearful, neutral-happy), c. Comparison of mean differences in drug effect for each emotion condition (collapsed across gain and loss conditions). *indicate significant differences.*

associated with punishment, but are expressed in areas associated with reward, this may indicate why oxytocin selectively affects learning during rewarding situations.

Lack of effect of oxytocin on punishment learning seemingly contradicts evidence that oxytocin can attenuate fear through oxytocinergic neurons in the central amygdala (Viviani et al., 2011; Knobloch et al., 2012) and imaging studies showing attenuation of amygdala response to fearful faces in normal subjects (Kirsch et al., 2005; Domes et al., 2007b) and patients with anxiety disorders (Labuschagne et al., 2010). However, as discussed, areas of the brain associated with punishment learning do not demonstrate a high expression of oxytocin receptors (Gimpl and Fahrenholz, 2001) and these areas of the brain are also not fundamental in the processing of fear. Therefore, although oxytocin can modulate fear through the amygdala (Viviani et al., 2011; Knobloch et al., 2012), it may not modulate the wider network associated with punishment (Guastella et al., 2008a).

More general negative effects of oxytocin have been found. For instance, although oxytocin has been shown to promote in-group cooperation, it also been shown to increase defensive aggression and derogatory behaviour toward out-group members (De Dreu et al., 2010; De Dreu et al., 2011). Furthermore, while literature supports

oxytocin's facilitatory effects on learning and memory for faces (Rimmele et al., 2009) and face expressions (Guastella et al., 2008b), there is evidence that suggests that oxytocin may impair memory functions in humans (Ferrier et al., 1980; Bruins et al., 1992; Heinrichs et al., 2004) and identity recognition of happy faces (Savaskan et al., 2008). However, this is the first study to demonstrate that oxytocin may be negatively modulating the effect of emotion on reward learning. In addition to this, we also observed that post-oxytocin participants rated those actors previously presented as happy with an 80% contingency as less attractive. It may well be that this effect is only noted in trials appointed an 80% contingency as they are indicative of more salient trials and provide the most information (i.e. actors are more likely to indicate a reward or punishment). Furthermore, as participants tend to get happy trials incorrect, they are more likely to receive a more unpleasant outcome. This may indicate why more salient happy actors are selectively rated as less attractive. Therefore, not only is oxytocin having detrimental learning effects, it is also causing unfavourable judgements of attractiveness after the event.

This study is an extension of a previous study, which examined the effects of oxytocin on associative learning, and demonstrated a reduction in aversion to angry

faces following oxytocin when financial feedback strongly favoured doing so (Evans et al., 2010). The latter study, however, had a number of limitations that the current study addressed. We introduced a neutral condition and also specifically examined any differences in the effects of oxytocin on reward and punishment learning. In contrast to observing a reduction in aversion to socially threatening stimuli, we found decreased learning specifically for happy trials in rewarding, but not punishing conditions. This suggests that results found in the associative learning task (Evans et al., 2010) may have been offset by reduced performance for happy faces. Moreover, by including neutral conditions, we were also able to look at how oxytocin may modulate the effect of emotion on social reward and punishment learning. We conclude that oxytocin, by affecting processing or the appraisal of happy faces, has a negative effect on a person's ability to efficiently learn rewards from happy faces.

Even though there were no differences shown in performance between the 'loss' and 'gain' blocks, we do report that response times appear to be significantly slower in the 'loss' compared to the 'gain' block. This may disguise any effects of learning from happy faces during this block and this could be a potential confounder when interpreting the findings. Furthermore, it could be argued that participants are actually learning very well and switching their responses, so that they are optimising their financial gains by chance. However, if one observes differences in the amount of money won for each condition (see Supplementary Information S3), analysis of the 'gain' block indicated that participants were also winning less for happy faces compared to neutral faces across both drug sessions. This suggests that participants are not only demonstrating reduced learning, but also reduced winnings for happy faces. Therefore, participants appear not be switching their responses to obtain the best outcome by chance.

In conclusion, this study demonstrates that oxytocin affects the processing of happy faces in certain contexts and this negatively impacts a person's ability to efficiently learn rewards from happy faces and also prompts adverse judgements of attractiveness after the event. This demonstrates how oxytocin detrimentally modulates the effects of emotion on social reward, but not punishment learning. Observing detrimental effects of oxytocin on emotional-cognitive processing in typical volunteers also indicates caution when studying potential therapeutic effects of oxytocin in clinical populations. Indeed oxytocin has been shown to increase empathy in people with autism spectrum conditions (Guastella et al., 2010) and has also been identified as a potential treatment for social anxiety disorder (Guastella et al., 2009). However, oxytocin detrimentally affects people with borderline personality disorder (Bartz et al., 2011). More research is needed to fully investigate the therapeutic effects of oxytocin in such clinical groups. Furthermore, we noted effects after one administration of the spray. More research is also needed to ascertain any long-term effects of utilising oxytocin to increase social functioning.

Supplementary material

For supplementary material accompanying this paper, visit http://dx.doi.org/10.1017/S1461145713001120

References

Andari E, Duhamel JR, Zalla T, Herbrecht E, Leboyer M, Sirigu A (2010) Promoting social behaviour with oxytocin in high-functioning autism spectrum disorders. Proc Natl Acad Sci USA 107:4389–4394.

Bartz J, Simeon D, Hamilton H, Kim S, Crystal S, Braun A, Vicens V, Hollander E (2011) Oxytocin can hinder trust and cooperation in borderline personality disorder. Soc Cogn Affect Neurosci 6:556–563.

Beck A, Steer R, Brown G (1996) Manual for Beck Depression Inventory-II. San Antonio, Tex: Psychological Corporation.

Bond A, Lader M (1974) The use of analogue scales in rating subjective feelings. Br J Med Psychol 47:211–217.

Born J, Lange T, Kern W, McGregor GP, Bickel U, Fehm HL (2002) Sniffing neuropeptides: a transnasal approach to the human brain. Nat Neurosci 5:514–516.

Bruins J, Hijman R, Van Ree JM (1992) Effect of a single dose of des-glycinamide-[Arg8]vasopressin or oxytocin on cognitive processes in young healthy subjects. Peptides 13:461–468.

Buijs RM, De Vries GJ, Van Leeuwen FW (1985) Oxytocin: Clinical and Laboratory Studies. In: The distribution and synaptic release of oxytocin in the central nervous system (Amico JA, Robinson AG, ed), pp 77–86. Amsterdam: BV Elsevier Science Publishers.

Calder AJ, Young AW (2005) Understanding the recognition of facial identity and facial expression. Nat Rev Neurosci 6:641–651.

D'Argembeau A, Van der Linden M (2011) Influence of facial expression on memory for facial identity: effects of visual features or emotional meaning? Emotion 11:199–202.

De Dreu CK, Greer LL, Handgraaf MJ, Shalvi S, Van Kleef GA, Baas M, Ten Velden FS, Van Dijk E, Feith SW (2010) The neuropeptide oxytocin regulates parochial altruism in inter-group conflict among humans. Science 328:1408–1411.

De Dreu CK, Greer LL, Van Kleef GA, Shalvi S, Handgraaf MJ (2011) Oxytocin promotes human ethnocentrism. Proc Natl Acad Sci U S A 108:1262–1266.

Di Simplicio M, Massey-Chase R, Cowen PJ, Harmer CJ (2009) Oxytocin enhances processing of positive *vs.* negative emotional information in healthy male volunteers. J Psychopharmacol 23:241–248.

Domes G, Heinrichs M, Michel A, Berger C, Herpertz SC (2007a) Oxytocin improves "mind-reading" in humans. Biol Psychiatry 61:731–733.

Domes G, Heinrichs M, Glascher J, Buchel C, Braus DF, Herpertz SC (2007b) Oxytocin attenuates amygdala responses to emotional faces regardless of valence. Biol Psychiatry 62:1187–1190.

Evans S, Shergill SS, Averbeck BB (2010) Oxytocin decreases aversion to angry faces in an associative learning task. Neuropsychopharmacology 35:2502–2509.

Ferrier BM, Kennett DJ, Devlin MC (1980) Influence of oxytocin on human memory processes. Life Sci 27:2311–2317.

Gamer M, Zurowski B, Buchel C (2010) Different amygdala subregions mediate valence-related and attentional effects of oxytocin in humans. Proc Natl Acad Sci U S A 107:9400–9405.

Gimpl G, Fahrenholz F (2001) The oxytocin receptor system: structure, function, and regulation. Physiol Rev 81:629–683.

Guastella AJ, Mitchell PB, Dadds MR (2008a) Oxytocin increases gaze to the eye region of human faces. Biol Psychiatry 63:3–5.

Guastella AJ, Mitchell PB, Mathews F (2008b) Oxytocin enhances the encoding of positive social memories in humans. Biol Psychiatry 64:256–258.

Guastella AJ, Howard AL, Dadds MR, Mitchell P, Carson DS (2009) A randomized controlled trial of intranasal oxytocin as an adjunct to exposure therapy for social anxiety disorder. Psychoneuroendocrinology 34:917–923.

Guastella AJ, Einfeld SL, Gray KM, Rinehart NJ, Tonge BJ, Lambert TJ, Hickie IB (2010) Intranasal oxytocin improves emotion recognition for youth with autism spectrum disorders. Biol Psychiatry 67:692–694.

Heinrichs M, Meinlschmidt G, Wippich W, Ehlert U, Hellhammer DH (2004) Selective amnesic effects of oxytocin on human memory. Physiol Behav 83:31–38.

Huber D, Veinante P, Stoop R (2005) Vasopressin and oxytocin excite distinct neuronal populations in the central amygdala. Science 308:245–248.

Hurlemann R, Patin A, Onur OA, Cohen MX, Baumgartner T, Metzler S, Dziobek I, Gallinat J, Wagner M, Maier W, Kendrick KM (2010) Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. J Neurosci 30:4999–5007.

Insel TR, Hulihan TJ (1995) A gender-specific mechanism for pair bonding: oxytocin and partner preference formation in monogamous voles. Behav Neurosci 109:782–789.

Kirsch P, Esslinger C, Chen Q, Mier D, Lis S, Siddhanti S, Gruppe H, Mattay VS, Gallhofer B, Meyer-Lindenberg A (2005) Oxytocin modulates neural circuitry for social cognition and fear in humans. J Neurosci 25:11489–11493.

Knobloch HS, Charlet A, Hoffmann LC, Eliava M, Khrulev S, Cetin AH, Osten P, Schwarz MK, Seeburg PH, Stoop R, Grinevich V (2012) Evoked axonal oxytocin release in the central amygdala attenuates fear response. Neuron 73:553–566.

Kosfeld M, Heinrichs M, Zak PJ, Fischbacher U, Fehr E (2005) Oxytocin increases trust in humans. Nature 435:673–676.

Labuschagne I, Phan KL, Wood A, Angstadt M, Chua P, Heinrichs M, Stout JC, Nathan PJ (2010) Oxytocin attenuates amygdala reactivity to fear in generalized social anxiety disorder. Neuropsychopharmacology 35:2403–2413.

Lee HJ, Macbeth AH, Pagani JH, Young WS III (2009) Oxytocin: the great facilitator of life. Prog Neurobiol 88:127–151.

Liu Y, Wang ZX (2003) Nucleus accumbens oxytocin and dopamine interact to regulate pair bond formation in female prairie voles. Neuroscience 121:537–544.

Lundqvist D, Flykt A, Öhman A (1998) The Karolinska Directed Emotional Faces – KDEF. Solna, Sweden: CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet.

Marsh AA, Yu HH, Pine DS, Blair RJ (2010) Oxytocin improves specific recognition of positive facial expressions. Psychopharmacology (Berl) 209:225–232.

Mattfeld AT, Gluck MA, Stark CEL (2011) Functional specialization within the striatum along both the dorsal/ventral and anterior/posterior axes during associative learning via reward and punishment. Learn Mem 18:703–711.

Monk CS, Nelson EE, McClure EB, Mogg K, Bradley BP, Leibenluft E, Blair RJ, Chen G, Charney DS, Ernst M, Pine DS (2006) Ventrolateral prefrontal cortex activation and attentional bias in response to angry faces in adolescents with generalized anxiety disorder. Am J Psychiatry 163:1091–1097.

Palminteri S, Justo D, Jauffret C, Pavlicek B, Dauta A, Delmaire C, Czernecki V, Karachi C, Capelle L, Durr A, Pessiglione M (2012) Critical roles for anterior insula and dorsal striatum in punishment-based avoidance learning. Neuron 76:998–1009.

Petrovic P, Kalisch R, Singer T, Dolan RJ (2008) Oxytocin attenuates affective evaluations of conditioned faces and amygdala activity. J Neurosci 28:6607–6615.

Popik P, Vos PE, Van Ree JM (1992) Neurohypophyseal hormone receptors in the septum are implicated in social recognition in the rat. Behav Pharmacol 3:351–358.

Rimmele U, Hediger K, Heinrichs M, Klaver P (2009) Oxytocin makes a face in memory familiar. J Neurosci 29:38–42.

Savaskan E, Ehrhardt R, Schulz A, Walter M, Schachinger H (2008) Post-learning intranasal oxytocin modulates human memory for facial identity. Psychoneuroendocrinology 33:368–374.

Schulze L, Lischke A, Greif J, Herpertz SC, Heinrichs M, Domes G (2011) Oxytocin increases recognition of masked emotional faces. Psychoneuroendocrinology 36:1378–1382.

Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC (1998) The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 59(Suppl. 20):22–33;quiz 34–57.

Spielberg CD, Gorsuch RL, Lushene RE (1983) Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologist Press.

Theodoridou A, Rowe AC, Penton-Voak IS, Rogers PJ (2009) Oxytocin and social perception: oxytocin increases perceived facial trustworthiness and attractiveness. Horm Behav 56:128–132.

Viviani D, Charlet A, van den Burg E, Robinet C, Hurni N, Abatis M, Magara F, Stoop R (2011) Oxytocin selectively gates fear responses through distinct outputs from the central amygdala. Science 333:104–107.

Witt DM, Winslow JT, Insel TR (1992) Enhanced social interactions in rats following chronic, centrally infused oxytocin. Pharmacol Biochem Behav 43:855–861.

Young LJ, Lim MM, Gingrich B, Insel TR (2001) Cellular mechanisms of social attachment. Horm Behav 40:133–138.

Zak PJ, Stanton AA, Ahmadi S (2007) Oxytocin increases generosity in humans. PLoS ONE 2:e1128.