

# Oxytocin can hinder trust and cooperation in borderline personality disorder

Jennifer Bartz,<sup>1</sup> Daphne Simeon,<sup>2</sup> Holly Hamilton,<sup>1</sup> Suah Kim,<sup>1</sup> Sarah Crystal,<sup>1</sup> Ashley Braun,<sup>1</sup> Victor Vicens,<sup>1</sup> and Eric Hollander<sup>3</sup>

<sup>1</sup>Department of Psychiatry, Mount Sinai School of Medicine, New York, NY, 10029, USA, <sup>2</sup>Beth Israel Medical Center and Albert Einstein College of Medicine, New York, NY, 10003, USA, and <sup>3</sup>Montefiore Medical Center University Hospital, Albert Einstein College of Medicine, New York, NY, 10467, USA

**We investigated the effects of intranasal oxytocin (OXT) on trust and cooperation in borderline personality disorder (BPD), a disorder marked by interpersonal instability and difficulties with cooperation. Although studies in healthy adults show that intranasal OXT increases trust, individuals with BPD may show an altered response to exogenous OXT because the effects of OXT on trust and pro-social behavior may vary depending on the relationship representations and expectations people possess and/or altered OXT system functioning in BPD. BPD and control participants received intranasal OXT and played a social dilemma game with a partner. Results showed that OXT produced divergent effects in BPD participants, decreasing trust and the likelihood of cooperative responses. Additional analyses focusing on individual differences in attachment anxiety and avoidance across BPD and control participants indicate that these divergent effects were driven by the anxiously attached, rejection-sensitive participants. These data suggest that OXT does not uniformly facilitate trust and pro-social behavior in humans; indeed, OXT may impede trust and pro-social behavior depending on chronic interpersonal insecurities, and/or possible neurochemical differences in the OXT system. Although popularly dubbed the ‘hormone of love’, these data suggest a more circumspect answer to the question of who will benefit from OXT.**

**Keywords:** oxytocin; trust; cooperation; social dilemma; borderline personality disorder; adult attachment

## INTRODUCTION

The ability to engage in pro-social, cooperative behavior is essential for developing and maintaining stable relationships. The neuropeptide oxytocin (OXT) has been shown to play a central role in pro-social behavior (Carter *et al.*, 1992; Panksepp, 1992; Carter, 1998; Insel and Young, 2001). Specifically, research in animals indicates that OXT is critically involved in pair-bond formation, separation distress and other aspects of attachment and affiliation (Lim and Young, 2006). Some evidence suggests that OXT is also involved in human pro-social behavior. For example, OXT, administered nasally, increased trusting behavior in a social dilemma game (Kosfeld *et al.*, 2005) and increased the perceived trustworthiness of faces (Theodoridou *et al.*, 2009). OXT may thus be a useful agent to increase pro-social behavior in individuals who have difficulties with such behaviors.

Borderline personality disorder (BPD) is characterized by affective instability, impulsivity—including impulsive aggression—and identity confusion (American Psychiatric Association, 2000). Interpersonal instability is also a core feature of BPD: individuals with BPD tend to have intense relationships marked by desperate attempts to avoid abandonment. Ironically, these reassurance-seeking strategies are often accompanied by efforts to downplay the importance of closeness and/or aggressive acts aimed at punishing significant others (Gunderson, 1996), leading to relationships marked by frequent arguments, repeated breakups and overall emotional volatility (Lieb *et al.*, 2004). Although interpersonal difficulties are most apparent in established relationships, individuals with BPD were shown to have difficulty sustaining cooperation in a social dilemma game played with a stranger (King-Casas *et al.*, 2008), suggesting that their difficulties can extend beyond their existing relationships. Given that OXT promotes pro-social, trusting behavior in healthy adults, OXT may be helpful in facilitating such behaviors in BPD.

Other research, however, suggests that the effects of OXT may differ in those with BPD. First, a recent study found that OXT (*vs* placebo) increased negative social emotions like envy and ‘schadenfreude’ (Shamay-Tsoory *et al.*, 2009); it was suggested that rather than having broad positive effects on social perception and behavior, OXT may increase the salience of social cues, thereby triggering the positive *or* negative

Received 26 February 2010; Accepted 8 September 2010

Advance Access publication 29 November 2010

We thank Geraldine Downey and Mark Baldwin for commenting on an earlier version of this article and Faby Gagne for advice on statistical analyses.

This research was made possible by Grant Number M01-RR00071 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NCRR or NIH. Jennifer Bartz is also funded by the Seaver Foundation and the National Institutes of Health to investigate the effects of intranasal oxytocin on social cognition in autism spectrum disorders.

Correspondence should be addressed to Jennifer Bartz, Department of Psychiatry, Mount Sinai School of Medicine, One Gustave L. Levy Place, Box 1230, New York, NY 10029, USA. E-mail: jennifer.bartz@mssm.edu.

emotions associated with them. Others have suggested that OXT increases approach behaviors (Kemp and Guastella, 2010), or affiliative drive more specifically (Taylor *et al.*, 2006). These alternate explanations of OXT function all suggest that individual differences—especially differences in the relationship representations and expectations people possess—and/or situational factors may critically moderate the effects of OXT on social perception and behavior. With respect to BPD, if OXT increases the salience of social cues, or increases affiliative drive, OXT may in fact exacerbate BPD individuals' chronic concerns about abandonment and trust, and their difficulties with cooperation, especially when administered in situations where such issues are salient.

Second, the OXT system may be dysregulated in BPD and, for this reason, may produce a differential response to exogenous OXT. Research shows that negative interpersonal experiences can impact the endogenous OXT system: lower cerebrospinal fluid (CSF) OXT levels have been observed in nursery- vs mother-reared monkeys (Winslow *et al.*, 2003), and in women who experienced childhood abuse and/or neglect (Heim *et al.*, 2008), and higher (in contrast to CSF) plasma OXT levels have been associated with self-reports of relationship distress (Taylor *et al.*, 2006), anxiety over relationships (Turner *et al.*, 1999) and social anxiety symptom severity in social phobia (Hoge *et al.*, 2008). Moreover, there is preliminary evidence that such socially 'at risk' individuals may differentially respond to exogenous OXT (Meinlschmidt and Heim, 2007). Given the centrality of interpersonal dysfunction in BPD, these individuals may be good candidates for neurochemical alterations within the OXT system (also see Stanley and Siever, 2009) and, thus, differential responsiveness to intranasal OXT.

### The present investigation

In this study, we investigated the effects of intranasal OXT on trust and cooperative behavior in healthy adults and adults with BPD. We tested two competing hypotheses about OXT function. On the one hand, studies in healthy adults suggest that intranasal OXT should facilitate trust and cooperation in both healthy control and BPD participants. On the other hand, individuals with BPD may show an altered response to intranasal OXT because the effects of OXT on trust and pro-social behavior vary as a function of the relationship representations one possesses and/or because of possible neurobiological differences in the OXT system in BPD.

In addition to investigating OXT response as a function of diagnostic status, we investigated whether individual differences in attachment anxiety and attachment avoidance moderate the effects of OXT on trust and pro-social behavior. Our rationale for this more nuanced approach was two-fold. First, individual differences in attachment anxiety and avoidance are important moderators of pro-social behavior. Avoidance is negatively associated with empathy (Mikulincer *et al.*, 2001), compassion and willingness to

help (Mikulincer *et al.*, 2005), volunteering (Gillath *et al.*, 2005), adhering to the communal script (Bartz and Lydon, 2008), identifying with values like benevolence and universalism (Mikulincer *et al.*, 2003) and cooperative helping on group tasks (Rom and Mikulincer, 2003). Attachment anxiety, by comparison, is associated with more ambivalent pro-social behavior, and data suggests a desire to affiliate that is sometimes hindered by interpersonal anxiety (e.g. Mikulincer *et al.*, 2001; Rom and Mikulincer, 2003; Bartz and Lydon, 2006). Second, several studies have shown considerable attachment heterogeneity in BPD, with some BPD individuals showing high levels of both anxiety and avoidance (i.e. 'fearful' types), and others showing high levels of anxiety only (i.e. 'preoccupied' types) (Agrawal *et al.*, 2004; Levy *et al.*, 2005; Aaronson *et al.*, 2006). If attachment anxiety and avoidance differentially affect pro-social behavior, failing to account for attachment heterogeneity could obscure important differences in the effects of OXT on pro-social behavior, especially in BPD participants.

## METHODS

### Participants

Thirteen healthy (7 males) and 14 adults with BPD (4 males) participated in this study; mean age was  $35 \pm 8$  years (BPD and control participants did not differ in age,  $t < 0.5$ , nor with respect to sex distribution,  $\chi^2(1, N = 27) = 1.78$ , *ns*). BPD participants were required to meet Diagnostic and Statistical Manual of Mental Disorders (DSM) IV-TR (American Psychiatric Association, 2000) BPD criteria; controls were excluded if they had any lifetime Axis I or II disorders. Additional exclusion criteria included: no psychotropic or other medications for at least 2 weeks prior to the study (5 weeks for fluoxetine); no current substance use disorder, major depression or eating disorders (anorexia or bulimia); no lifetime schizophrenia or bipolar I disorder; no mental retardation and no medical or neurological illness. Pregnant, lactating or menopausal females were also excluded. The study was approved by the Mount Sinai School of Medicine Institutional Review Board and all participants gave informed consent prior to participation.

### Design and procedures

#### Overview

Participants were first evaluated by a study psychiatrist; diagnostic eligibility was confirmed using the Structured Clinical Interview for DSM-IV Axis I Disorders (First *et al.*, 2002) and Structured Interview for DSM-IV Personality Disorders SCID-II (First *et al.*, 1994). Eligible participants returned 1–2 weeks later for the OXT challenge.

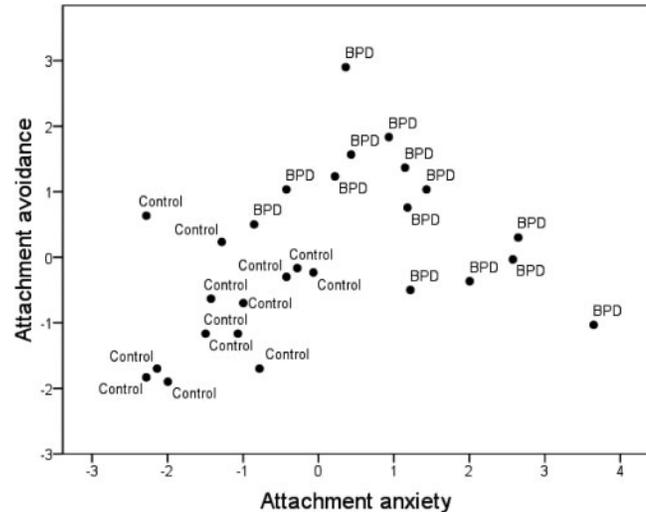
On the day of the challenge, participants completed some questionnaires, including the Experience in Close Relationship scale (ECR; Brennan *et al.*, 1998), which measures attachment anxiety and avoidance (see below). Baseline mood was assessed at this point with the Profile of Mood States (POMS; McNair *et al.*, 1992). Participants then

randomly received 40IU intranasal OXT (Syntocinon, Novartis) or placebo; 14 participants received OXT (six BPD) and 13 participants received placebo (eight BPD). Participants and experimenter were blind to drug condition. Approximately 35-min later, participants completed the POMS again to assess changes in mood as a function of drug; no mood changes were observed.<sup>1</sup> The experimenter then introduced participants to the Assurance Game (AG; Kollock, 1998), a variation of the classic Prisoner's Dilemma (PD) involving salient trust issues (see below). After this introduction, the experimenter confirmed that participants understood the AG and payoff matrix by giving them a brief quiz. Participants were then briefly introduced to their partner (a confederate) who, they were told, would be playing the AG with them from another room. After confirming there were no further questions, the experimenter told participants to begin the AG. To increase importance of their strategic choices, participants were told they could keep their earnings from the AG.

After the AG participants completed some additional tasks (not reported here), and were evaluated for side effects (with the exception of one participant reporting a headache, no side effects were reported as a function of OXT/placebo administration). Participants were paid \$75 in compensation, plus whatever they made from the social dilemma game. (This study was part of a two-day study; however, social dilemma task analyses were conducted on day 1 data only because of concerns about expectancy effects due to previous experiences of partner cooperation on the AG.)

### Experience in close relationship scale

The ECR (Brennan *et al.*, 1998) is a widely used and highly reliable self-report instrument for assessing attachment anxiety and avoidance in adults. The ECR consists of 36 items, 18 reflecting attachment anxiety (i.e. sensitivity to and anxiety about rejection/abandonment) and 18 reflecting attachment avoidance (i.e. discomfort with and desire to avoid closeness and intimacy).<sup>2</sup> Participants indicate on a 7-point scale the extent to which they agree/disagree with each item in terms of how they generally experience close relationships. Importantly, the ECR has been administered to adults with BPD, and data suggest that it is an appropriate instrument to assess attachment in this population (Levy *et al.*, 2005; Scott *et al.*, 2009). Not surprisingly, BPD participants were more anxiously attached ( $M=4.46$ ;  $SD=1.24$ ) than controls ( $M=2.01$ ;  $SD=0.76$ ),  $t(25)=6.15$ ,  $P<0.001$  and BPD participants were more avoidantly attached ( $M=4.32$ ;  $SD=1.04$ ) than controls ( $M=2.75$ ;  $SD=0.83$ ),  $t(25)=4.32$ ,  $P<0.001$ . Moreover, as expected, there was considerable attachment heterogeneity in the BPD group;



**Fig. 1** Scatter-plot displaying individual differences in attachment anxiety (mean-centered) and attachment avoidance (mean-centered), as measured by the ECR, in healthy control and borderline personality disorder (BPD) participants. Consistent with prior research, BPD participants fell into either the high anxious, low avoidant ('preoccupied') or high anxious, high avoidant ('fearful') quadrants of the ECR indicating the heterogeneous nature of attachment in BPD.

as depicted in Figure 1, BPD participants fell into either the high anxious, low avoidant ('preoccupied') or high anxious, high avoidant ('fearful') quadrants of the ECR.

### Assurance game

The Assurance Game (AG; Kollock, 1998) is a variation of the PD, involving salient trust issues. Specifically, whereas the PD pulls for self-interest by allotting the highest payoff for defection, the AG locates the self-interested and interpersonal solution in the same, mutual cooperation cell (i.e. both players make the most money—\$6 each—when they both cooperate). However, each player should only cooperate if he/she trusts that the other player will cooperate. If a player is mistrustful, he/she should defect, which is sub-optimal because the player only makes \$2, which is less than he/she would have made in the mutual cooperation scenario, but it is preferable to risking partner defection and making \$0 (see Figure 2 for AG payoff matrix).

Participants played three consecutive rounds of the AG in which they received cooperative feedback from their partner (we programmed the computer to make the partner cooperate on all three rounds, which allowed us to aggregate participants' responses across rounds for a more reliable index of trust and pro-social behavior). On each round, participants indicated their strategic choice by typing 1 for Strategy A, or 2 for Strategy B (2s were re-coded as 0s for data analyses), which reflect cooperation and defection, respectively (although the words 'cooperate' and 'defect' were never used during the testing session). Participants also indicated which strategy they thought their partner chose on that round using a 5-point scale (1 = Definitely Strategy A,

<sup>1</sup>Analyses of variance looking at the effect of OXT on mood change (i.e. 30-min post-drug mood minus baseline mood) showed no effects of OXT on any of the POMS mood change indices, all  $ps>0.1$ , nor were there any OXT by group interactions, all  $ps>0.1$ .

<sup>2</sup>We used a 29-item version of the ECR, which was modified to reduce item redundancy (Collins, 2005, personal communication).

Assurance Game Payoff Matrix

Partner	Participant	
	Strategy A (cooperate)	Strategy B (defect)
Strategy A (cooperate)	You get \$6 Your partner gets \$6	You get \$4 Your partner gets \$0
Strategy B (defect)	You get \$0 Your partner gets \$4	You get \$2 Your partner gets \$2

**Fig. 2** Assurance Game payoff matrix. Participants' payoff for each round is a function of their strategic choice and their partner's strategic choice. As depicted, participants make the most money (\$6) when they and their partner choose strategy A; however, participants should only choose strategy A if they trust their partner will also choose strategy A. The decision involves an element of trust because participants must make their strategic choice before they know what choice their partner made on that round. (Note: to circumvent socially desirable responses, the words 'cooperate' and 'defect' were never used; participants only saw 'Strategy A' and 'Strategy B').

2 = Probably Strategy A, 3 = 50–50 Strategy A or Strategy B, 4 = Probably Strategy B and 5 = Definitely Strategy B; this scale was reverse scored for data analyses so that higher numbers reflect more cooperative expectations). This was our index of partner expectations, or trust. Finally, participants indicated what strategy they would have chosen on that round 'If they knew that their partner would choose Strategy A' (i.e. if they knew their partner would cooperate) using the same 5-point scale (again, the scale was reverse scored for data analyses so higher numbers reflect more cooperative responses). This hypothetical scenario was included to distinguish self-protective from aggressive/hostile motives underlying defection. Specifically, participants can defect in the AG for two reasons. As noted, they can defect to protect themselves against partner defection (Figure 2, lower right matrix cell). However, if they know their partner is going to cooperate, there should be no reason to protect against partner defection. Thus, the strategic choice to defect in the hypothetical partner cooperation scenario can be considered a hostile strategic response since the other player makes \$0 compared to the \$6 he/she would have made had the participant cooperated (Figure 2, upper right matrix cell).

The AG was administered on a laptop computer and the payoff matrix was displayed at all times as a reminder. Participants were informed that they and their partner would make their strategic choices independently and simultaneously so that neither would know the strategic choice of the other when making their choice; however, after each round of the AG, participants were told the outcome for that round and, thus, could infer their partner's strategy for the previous round (this disclosure also served to make participants more accountable for their strategic choice). Participants answered the partner strategy expectation and strategic response to partner hypothetical cooperation questions before finding out what strategy their partner actually chose on each round.

### Statistical analyses

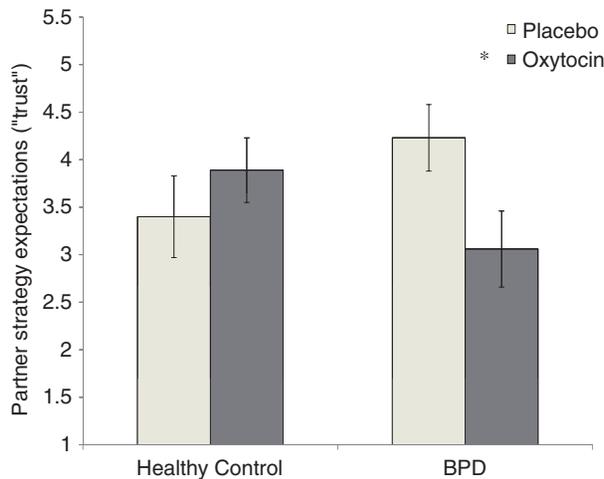
As noted, we adopted two data analytic strategies. We first used analysis of variance (ANOVA) to investigate the effects of group (BPD *vs* healthy control) and drug (OXT *vs* placebo) on mean trust, strategic response to partner hypothetical cooperation and cooperative behavior across rounds 1–3 of the AG game; age was included as a covariate because of considerable age variability (range 23–53 years) in this sample. These analyses were followed by t-tests to probe specific group differences. We then adopted a more nuanced approach, collapsing across diagnostic categories to look at whether individual differences in attachment anxiety and attachment avoidance moderate the effects of OXT on trust and pro-social behavior. Specifically, we conducted regression analyses on all 27 participants to look at the effects of drug (dummy coded: 1 = OXT and 0 = placebo), and mean-centered attachment anxiety and mean-centered avoidance (entered in step 1), and their two- and three-way interactions (entered in steps 2 and 3, respectively) on mean trust, response to hypothetical cooperation, and cooperative behavior; again, age was included as a covariate. Regression analyses were followed by simple slope analyses comparing each dependent variable in the OXT and placebo conditions at one standard deviation below and above the means for attachment anxiety and attachment avoidance. All analyses were tested at  $P < .05$  (two-tailed).<sup>3</sup>

## RESULTS

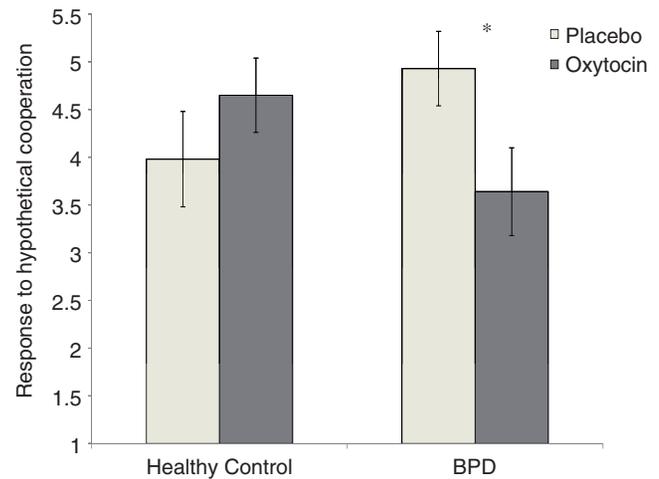
### The effects of group (BPD *vs* control) and OXT on trust and cooperation

ANOVA revealed a significant group (BPD *vs* healthy control)  $\times$  drug (OXT *vs* placebo) interaction for trust,

<sup>3</sup>Because less is known about the effects of exogenous OXT in women, we also ran all analyses with sex in the model; there were no significant main effects of sex, and sex did not significantly moderate any of the effects reported here (all  $ps > 0.1$ ).



**Fig. 3** Partner strategy expectations ('trust') following oxytocin or placebo for healthy control and BPD participants. Higher numbers reflect more cooperative partner expectations, or greater trust. Error bars represent standard errors of the difference between means. \* $P < 0.05$  (two-tailed).



**Fig. 4** Strategic response to partner hypothetical cooperation following oxytocin or placebo for healthy control and BPD participants. Higher numbers reflect more cooperative strategic choices. Error bars represent standard errors of the difference between means. \* $P < 0.05$  (two-tailed).

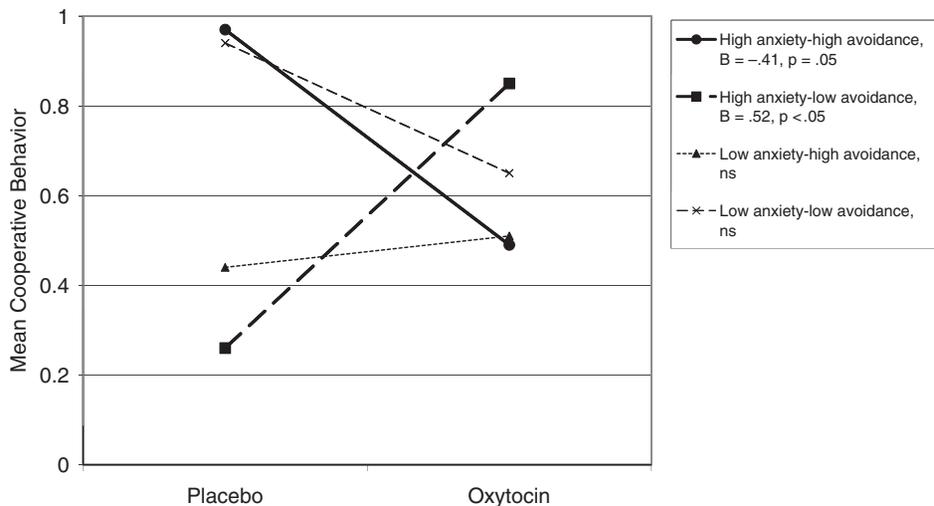
$F(1, 22) = 4.83$ ,  $P < 0.05$ , and for response to partner hypothetical cooperation,  $F(1, 22) = 5.06$ ,  $P < 0.05$ . As depicted in Figures 3 and 4, BPD participants expected their partner to be significantly less cooperative following OXT ( $M = 3.06$ ;  $SE = 0.40$ ) vs placebo ( $M = 4.23$ ;  $SE = 0.34$ ),  $t(22) = -2.25$ ,  $P < 0.05$  (Figure 3), and they were significantly more likely to defect in response to partner hypothetical cooperation following OXT ( $M = 3.64$ ;  $SE = 0.46$ ) vs placebo ( $M = 4.93$ ;  $SD = 0.40$ ),  $t(22) = -2.17$ ,  $P < 0.05$  (Figure 4). Healthy controls, by comparison, showed higher trusting expectations following OXT ( $M = 3.89$ ;  $SE = 0.34$ ) vs placebo ( $M = 3.39$ ;  $SE = 0.43$ ), and were more cooperative in the hypothetical cooperation scenario following OXT ( $M = 4.65$ ;  $SE = 0.39$ ) vs placebo ( $M = 3.98$ ;  $SE = 0.50$ ), but neither of these effects reached statistical significance (both  $t < 1.5$ ,  $ns$ ). Finally, ANOVAs revealed no effects of group or drug for actual cooperative behavior (all  $F_s < 1$ ).

### The effects of attachment anxiety, attachment avoidance and OXT on trust and cooperation

Regression analyses replicate and extend these findings. Consistent with the ANOVAs, regression analyses revealed a significant OXT  $\times$  attachment anxiety interaction for trust,  $B = -0.62$ ,  $t(19) = -2.26$ ,  $P < 0.05$ . Simple slope analyses indicate that OXT resulted in significantly less trusting expectations for anxiously attached, rejection-sensitive participants,  $B = -1.36$ ,  $t(22) = -2.36$ ,  $P < 0.05$ , whereas less anxiously attached participants showed no difference in trusting expectations in the OXT versus placebo conditions,  $t < 1$ ,  $ns$  (although the slope was in the predicted direction,  $B = 0.42$ , with trust increasing in the OXT condition for controls). Main effects and other two-way interactions were not significant.

Also paralleling the group analyses, regression analyses revealed a significant effect of attachment anxiety,  $B = 0.43$ ,  $t(19) = -2.32$ ,  $P < 0.05$ , which was qualified by a significant OXT  $\times$  attachment anxiety interaction for response to partner hypothetical cooperation,  $B = -0.66$ ,  $t(19) = -2.22$ ,  $P < 0.05$ . Simple slope analyses indicate that OXT resulted in significantly less cooperation in the hypothetical scenario for anxiously attached, rejection-sensitive participants,  $B = -1.69$ ,  $t(22) = -2.72$ ,  $P = 0.012$ , whereas less anxiously attached participants showed no difference in cooperation in the OXT versus placebo conditions,  $t < 1.55$ ,  $ns$  (again, though, the slope was in the predicted direction,  $B = 0.93$ , with cooperation increasing in the OXT condition for controls). Other main effects and other two-way interactions were not significant.

Finally, with respect to actual cooperative behavior, regression analyses revealed a significant attachment anxiety  $\times$  avoidance interaction,  $B = 0.12$ ,  $t(18) = -2.51$ ,  $P < 0.05$ , which was qualified by a significant OXT  $\times$  attachment anxiety  $\times$  attachment avoidance interaction,  $B = -0.18$ ,  $t(18) = -2.59$ ,  $P = 0.018$ . Simple slope analyses indicate that OXT primarily affected anxiously attached, rejection-sensitive participants, but the direction of the effect depended on avoidance, with OXT increasing actual cooperative behavior for anxiously attached/low avoidant individuals,  $B = 0.52$ ,  $t(18) = 2.24$ ,  $P < 0.05$ , but decreasing actual cooperative behavior for anxiously attached/high avoidant individuals,  $B = -0.41$ ,  $t(18) = -2.09$ ,  $P = 0.05$  (Figure 5). Simple slope analyses for the low anxious/low avoidant ('secure'), and low anxious/high avoidant ('dismissive') participants showed no differences in actual cooperative behavior in the OXT vs placebo conditions, both  $t < 1.65$ ,  $ns$ . Main effects and other two-way interactions were not significant.



**Fig. 5** Actual cooperative behavior for all (BPD and control) participants as a function of OXT or placebo, and individual differences in attachment anxiety and attachment avoidance. Higher numbers reflect more cooperative behavior. Simple slope analyses showed that OXT primarily modulated cooperative behavior for the anxiously attached participants, but the direction of the effect depended on avoidance level, with OXT significantly increasing cooperative behavior for high anxious, low avoidant ('preoccupied') participants ( $B = 0.52$ ,  $t(18) = 2.24$ ,  $P < 0.05$ ; bold dashed line), but significantly decreased cooperative behavior for high anxious, high avoidant ('fearful') participants ( $B = -0.41$ ,  $t(18) = -2.09$ ,  $P = 0.05$ ; bold solid line). Low anxious participants showed no change in cooperative behavior as a function of OXT (both  $p$ s  $> 0.1$ ).

## DISCUSSION

We aimed to investigate the effects of OXT—a neuropeptide implicated in attachment and pro-social behavior in animals and trust in healthy humans—on trust and pro-social behavior in individuals with BPD, who are characterized by interpersonal and affective instability, impulsive aggression, and difficulties sustaining cooperative, pro-social behavior. BPD participants had significantly less trusting expectations about their partner and were significantly more likely to defect in response to partner cooperation in a hypothetical scenario following intranasal OXT compared to placebo. Recall that in the hypothetical cooperation scenario, participants were asked what strategy they would choose 'if they knew their partner would cooperate'; thus, choosing to defect is, arguably, a hostile strategic response since the other player makes nothing (compared to the \$6 the other player would have made had the participant cooperated); it is also an irrational strategy since it costs participants monetarily (participants make \$4 rather than \$6 they would have made had they cooperated). Self-protective motives can likely be ruled out here because there should be no need to self-protect when you know your partner will cooperate. Thus it appears that the BPD participants were guided less by rational game norms and more by an interpersonal desire to punish their partner following OXT. In contrast to previous studies (Kosfeld *et al.*, 2005; Theodoridou *et al.*, 2009), we did not observe a significant increase in trust and/or pro-social behavior in our healthy controls participants; however, this may have been due to ceiling effects on the AG.

Additional analyses focusing on individual differences in attachment anxiety and avoidance across BPD and control participants replicate and extend the diagnostic-level

analyses. These analyses showed that the divergent effects of OXT on trust and strategic response to partner cooperation in the hypothetical scenario were primarily driven by the anxiously attached, rejection-sensitive participants. Moreover, whereas analyses based on diagnosis (BPD *vs* control) showed no effect of diagnosis or OXT on actual cooperative behavior, individual difference analyses revealed that OXT promoted actual cooperative behavior for anxiously attached but low avoidant (i.e. intimacy seekers) individuals but impeded cooperative behavior for anxiously attached, intimacy-avoidant individuals. The fact that we observed no differences in actual cooperative behavior when focusing on diagnostic category highlights the heterogeneity of BPD and the importance of considering individual differences in the motivation to approach versus avoid intimacy in this population, especially when looking at pro-social behavior.

There are a few possible explanations for the divergent effects of OXT observed in this study. First, as noted, rather than increasing positive social emotions like trust, OXT may increase the salience of social cues and, therefore, may trigger a range of emotions and behaviors—both positive and negative—involved in regulating social interactions (Shamay-Tsoory *et al.*, 2009). Thus, whether OXT promotes or hinders pro-social behavior depends on the individual and their social repertoire, and/or the social context. In this study, increasing the salience of social cues in the context of the Assurance Game (which was selected because it made trust issues salient) may have activated chronic concerns about trust and closeness in BPD/anxiously attached participants and they may have relied on their pre-existing but maladaptive strategies (i.e. defect/punish partner) to cope. A second related possibility is that OXT may activate

approach behaviors (Kemp and Guastella, 2010) and/or a desire to affiliate (Taylor, 2006) but, again, this motivation to affiliate may remind BPD/anxiously attached participants of previous experiences when affiliation has gone awry and set in motion their chronic concerns about trust and closeness. A third possibility is that the OXT system may be dysregulated in BPD. As noted, negative interpersonal experiences are associated with altered OXT levels measured in the CSF (Winslow et al., 2003; Heim et al., 2008), and plasma (Turner et al., 1999; Taylor et al., 2006; Hoge et al., 2008). It is plausible that such neurobiological differences in the OXT system may produce a differential response to exogenous OXT (e.g. Winslow and Insel, 1991; Taylor et al., 2006; Meinschmidt and Heim, 2007). Given the centrality of interpersonal dysfunction in BPD, these individuals would be good candidates for such neurochemical alterations within the OXT system. All of these hypotheses should be investigated in future research.

Although these data suggest that it might be difficult to alter chronic interpersonal expectancies and interaction patterns from a one-time administration of OXT, these findings do not necessarily argue against OXT therapy in BPD. If OXT enhances the salience of social cues and/or activates affiliative motives, it might be an ideal way to facilitate the learning of new relational schemas and interpersonal skills, for example, in combination with cognitive-behavioral (e.g. Guastella et al., 2009) or behavioral interventions.

In conclusion, these data suggest that OXT does not uniformly facilitate trust and pro-social behavior in humans; indeed, OXT may impede trust and pro-social behavior depending on diagnosis (BPD) and/or chronic interpersonal insecurities combined with situational factors that heighten those insecurities. Neurochemical alterations within the OXT system may also influence response to exogenous OXT (Taylor et al., 2006; Heim et al., 2008). Although popularly dubbed the 'hormone of love', these findings highlight the importance of considering individual differences in diagnosis and/or social motivation (approach/avoid intimacy) when evaluating the therapeutic potential of OXT for targeting social functioning deficits.

### Conflict of Interest

Dr Hollander has applied for a patent for oxytocin in autism and related conditions. Dr Bartz, Dr Simeon, Ms Hamilton, Ms Kim, Ms Crystal, Ms Braun and Dr Vicens declare no potential conflict of interest.

### REFERENCES

- Aaronson, C.J., Bender, D.S., Skodol, A.E., Gunderson, J.G. (2006). Comparison of attachment styles in borderline personality disorder and obsessive-compulsive personality disorder. *Psychiatric Quarterly*, 77, 69–80.
- Agrawal, H.R., Gunderson, J., Holmes, B.M., Lyons-Ruth, K. (2004). Attachment studies with borderline patients: a review. *Harvard Review of Psychiatry*, 12, 94–104.
- American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn, Washington, DC: American Psychiatric Association.
- Bartz, J.A., Lydon, J.E. (2006). Navigating the interdependence dilemma: attachment goals and the use of communal norms with potential close others. *Journal of Personality and Social Psychology*, 91, 77–96.
- Bartz, J.A., Lydon, J.E. (2008). Relationship-specific attachment, risk regulation, and communal norm adherence in close relationships. *Journal of Experimental Social Psychology*, 44, 655–63.
- Brennan, K.A., Clark, C.L., Shaver, P.R. (1998). Self-report measurement of adult attachment: an integrative overview. In: Simpson, J.A., Rholes, W.S., editors. *Attachment Theory and Close Relationships*. New York: Guilford Press, pp. 46–76.
- Carter, C.S. (1998). Neuroendocrine perspectives on social attachment and love. *Psychoneuroendocrinology*, 23, 779–818.
- Carter, C.S., Williams, J.R., Witt, D.M., Insel, T.R. (1992). Oxytocin and social bonding. *Annals of the New York Academy of Sciences*, 652, 204–11.
- First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W. (2002). *Structured Clinical Interview for DSM-IV-TR Axis I Disorders-Patient Edition (SCID-I/P, 11/2002 revision)*. New York: Biometrics Research Department, New York State Psychiatric Institute.
- First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W., Benjamin, L. (1994). *Structured Clinical Interview for DSM-IV-TR Axis II Personality Disorders (SCID-II), (Version 2.0)*. New York: Biometrics Research Department, New York State Psychiatric Institute.
- Gillath, O., Shaver, P.R., Mikulincer, M., Nitzberg, R.A., Erez, A., Van Ijzendoorn, M.H. (2005). Attachment, caregiving, and volunteerism: placing volunteerism in an attachment-theoretical framework. *Personal Relationships*, 12, 425–46.
- Guastella, A.J., Howard, A.L., Dadds, M.R., Mitchell, P., Carson, D.S. (2009). A randomized controlled trial of intranasal oxytocin as an adjunct to exposure therapy for social anxiety disorder. *Psychoneuroendocrinology*, 34, 917–23.
- Gunderson, J.G. (1996). The borderline patient's intolerance of aloneness: insecure attachments and therapist availability. *American Journal of Psychiatry*, 153, 752–8.
- Heim, C., Young, L.J., Newport, D.J., Mletzko, T., Miller, A.H., Nemeroff, C.B. (2009). Lower CSF oxytocin concentrations in women with a history of childhood abuse. *Molecular Psychiatry*, 14, 954–8.
- Hoge, E.A., Pollack, M.H., Kaufman, R.E., Zak, P.J., Simon, N.M. (2008). Oxytocin levels in social anxiety disorder. *CNS Neuroscience and Therapeutics*, 14, 165–70.
- Insel, T.R., Young, L.J. (2001). The neurobiology of attachment. *Nature Reviews Neuroscience*, 2, 129–36.
- Kemp, A.H., Guastella, A.J. (2010). Oxytocin: prosocial behavior, social salience, or approach-related behavior? *Biological Psychiatry*, 67, e33–4, author reply e35.
- King-Casas, B., Sharp, C., Lomax-Bream, L., Lohrenz, T., Fonagy, P., Montague, P.R. (2008). The rupture and repair of cooperation in borderline personality disorder. *Science*, 321, 806–10.
- Kollock, P. (1998). Social dilemmas: the anatomy of cooperation. *Annual Review of Sociology*, 24, 183–214.
- Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U., Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, 435, 673–6.
- Levy, K.N., Meehan, K.B., Weber, M., Reynoso, J., Clarkin, J.F. (2005). Attachment and borderline personality disorder: implications for psychotherapy. *Psychopathology*, 38, 64–74.
- Lieb, K., Zanarini, M.C., Schmahl, C., Linehan, M.M., Bohus, M. (2004). Borderline personality disorder. *Lancet*, 364, 453–61.
- Lim, M.M., Young, L.J. (2006). Neuropeptidergic regulation of affiliative behavior and social bonding in animals. *Hormones and Behavior*, 50, 506–17.
- McNair, D.M., Lorr, M., Droppelman, L.F. (1992). *EdiITS Manual for the Profile of Mood States*. San Diego, CA: Educational & Industrial Testing Service.

- Meinlschmidt, G., Heim, C. (2007). Sensitivity to intranasal oxytocin in adult men with early parental separation. *Biological Psychiatry*, *61*, 1109–11.
- Mikulincer, M., Shaver, P.R., Gillath, O., Nitzberg, R.A. (2005). Attachment, caregiving, and altruism: boosting attachment security increases compassion and helping. *Journal of Personality and Social Psychology*, *89*, 817–39.
- Mikulincer, M., Gillath, O., Halevy, V., Avihou, N., Avidan, S., Eshkoli, N. (2001). Attachment theory and reactions to others' needs: evidence that activation of the sense of attachment security promotes empathic responses. *Journal of Personality and Social Psychology*, *81*, 1205–24.
- Mikulincer, M., Gillath, O., Sapir-Lavid, Y., Yaakobi, E., Arias, K., Tal-Aloni, L., et al. (2003). Attachment theory and concern for others' welfare: evidence that activation of the sense of secure base promotes endorsement of self-transcendence values. *Basic and Applied Social Psychology*, *25*, 299–312.
- Panksepp, J. (1992). Oxytocin effects on emotional processes: separation distress, social bonding, and relationships to psychiatric disorders. *Annals of the New York Academy of Sciences*, *652*, 243–52.
- Rom, E., Mikulincer, M. (2003). Attachment theory and group processes: the association between attachment style and group-related representations, goals, memories, and functioning. *Journal of Personality and Social Psychology*, *84*, 1220–35.
- Scott, L.N., Levy, K.N., Pincus, A.L. (2009). Adult attachment, personality traits, and borderline personality disorder features in young adults. *Journal of Personality Disorders*, *23*, 258–80.
- Shamay-Tsoory, S.G., Fischer, M., Dvash, J., Harari, H., Perach-Bloom, N., Levkovitz, Y. (2009). Intranasal administration of oxytocin increases envy and schadenfreude (gloating). *Biological Psychiatry*, *66*, 864–70.
- Stanley, B., Siever, L.J. (2009). The interpersonal dimension of borderline personality disorder: toward a neuropeptide model. *American Journal of Psychiatry*, *167*, 24–39.
- Taylor, S.E. (2006). Tend and befriend: biobehavioral bases of affiliation under stress. *Current Directions in Psychological Science*, *15*, 273–7.
- Taylor, S.E., Gonzaga, G.C., Klein, L.C., Hu, P., Greendale, G.A., Seeman, T.E. (2006). Relation of oxytocin to psychological stress responses and hypothalamic-pituitary-adrenocortical axis activity in older women. *Psychosomatic Medicine*, *68*, 238–45.
- Theodoridou, A., Rowe, A.C., Penton-Voak, I.S., Rogers, P.J. (2009). Oxytocin and social perception: oxytocin increases perceived facial trustworthiness and attractiveness. *Hormones and Behavior*, *56*, 128–32.
- Turner, R.A., Altemus, M., Enos, T., Cooper, B., McGuinness, T. (1999). Preliminary research on plasma oxytocin in normal cycling women: investigating emotion and interpersonal distress. *Psychiatry*, *62*, 97–113.
- Winslow, J.T., Insel, T.R. (1991). Social status in pairs of male squirrel monkeys determines the behavioral response to central oxytocin administration. *The Journal of Neuroscience*, *11*, 2032–8.
- Winslow, J.T., Noble, P.L., Lyons, C.K., Sterk, S.M., Insel, T.R. (2003). Rearing effects on cerebrospinal fluid oxytocin concentration and social buffering in rhesus monkeys. *Neuropsychopharmacology*, *28*, 910–8.