Author's Accepted Manuscript

The social salience hypothesis of oxytocin
Simone G. Shamay-Tsoory, Ahmad Abu-Akel

PII: S0006-3223(15)00639-3
DOI: http://dx.doi.org/10.1016/j.biopsych.2015.07.020
Reference: BPS12628

To appear in: Biological Psychiatry

Cite this article as: Simone G. Shamay-Tsoory, Ahmad Abu-Akel, The social salience hypothesis of oxytocin, Biological Psychiatry, http://dx.doi.org/10.1016/j.biopsych.2015.07.020

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Title: The social salience hypothesis of oxytocin

Short title: The social salience hypothesis of oxytocin

Authors: Simone G. Shamay-Tsoory¹*, Ahmad Abu-Akel²

¹ Department of Psychology, University of Haifa, Haifa, Israel
² School of Psychology, University of Birmingham, U.K.

* Corresponding author: Simone G. Shamay-Tsoory, Department of Psychology, University of Haifa, Haifa 31905, Israel. Email: sshamay@psy.haifa.ac.il

Key words: oxytocin, social salience, dopamine, context, attention, amygdala

Abstract

Oxytocin is a nonapeptide that also serves as a neuromodulator in the human central nervous system. Over the last decade, a sizeable body of literature has examined its effects on social behavior in humans. These studies show that oxytocin modulates various aspects of social behaviors such as empathy, trust, ingroup preference and memory of socially-relevant cues. Several theoretical formulations have attempted to explain the effects of oxytocin. The prosocial account argues that oxytocin mainly enhances affiliative prosocial behaviors; the fear/stress theory suggests that oxytocin affects social performance by attenuating stress; and the in-/out-group approach proposes that oxytocin regulates cooperation and conflict among humans in the context of intergroup relations. Nonetheless, accumulating evidence reveals that the effects of
oxytocin are dependent on a variety of contextual aspects, the individual’s characteristics, and can induce antisocial effects including aggression and envy. In an attempt to reconcile these accounts, we suggest a theoretical framework that focuses on the overarching role of oxytocin in regulating the salience of social cues through its interaction with the dopaminergic system. Crucially, the salience effect modulates attention orienting responses to external contextual social cues (e.g., competitive vs. cooperative environment), but which is dependent on baseline individual differences such as gender, personality traits and degree of psychopathology. This view could have important implications for the therapeutic applications of oxytocin in conditions characterized with aberrant social behavior.

Introduction

Oxytocin and the Social Salience Hypothesis

Oxytocin (OT) is a nonapeptide, synthesized in the paraventricular nucleus and the supraoptic nucleus of the hypothalamus, and released by the pituitary gland. Animal studies show that OT has an established hormonal function in uterine contractions and lactating in nursing females (1), as well as a central role in social behavior and affiliation (2-6). The modulatory role of OT on animal’s social behaviors has led to strong interest in its effects on social behaviors in humans, showing largely similar effects [for a review see (7, 8)].

Several accounts have been proposed to explain the mechanism by which OT exerts its effects. These include the prosocial account which argues that OT mainly enhances affiliative
prosocial behaviors (9, 10); the fear/stress account which suggests that OT affects social performance by attenuating stress (11); and the in-/out-group approach which proposes that OT regulates cooperation and conflict among humans in the context of intergroup relations (12, 13). Yet, accumulating evidence reveals that the effects of OT are not always positive and may be context dependent.

In one of the initial studies that tested the possible non-prosocial effects of OT, it was found that following intranasal administration of OT (inOT) feelings of envy were remarkably increased when the other player gained more money, as were feelings of schadenfreude when the participant gained more money than the other player (14). In this study, the social salience hypothesis was introduced as a theoretical framework to understanding the effects of OT, and suggested that elevated envy and schadenfreude may have been the consequence of OT increasing the salience of competitive cues in the task (14, 15).

A growing list of studies have further demonstrated that OT can induce antisocial effects including aggression (14, 16) and that high levels of endogenous OT are associated with relational distress and interpersonal difficulties (17, 18). In addition, it was found that the effects of OT interact with contextual cues, gender, psychiatric conditions and personality traits (19-22). Building on this evidence, the social salience hypothesis of OT was further expanded by Bartz et al. (23) suggesting that the effects of OT are interactively constrained by features of situations and/or individuals, and by Olff et al. (24) who suggested that OT increases sensitivity to social cues depending on contextual variables and interindividual factors.

While previous discussions proposed that the mechanism underlying the effects of OT is related to its role in regulating the salience of social stimuli, the nature of this mechanism is unknown. Here we present converging evidence, from both animal and human studies, which
supports a theoretical framework that integrates the prosocial, stress and intergroup accounts by focusing on the overarching role of OT in increasing the salience of social cues. It should be noted that the data reported in inOT studies and those examining plasma OT levels (25), are correlational in nature and therefore the implications for central OT activity are controversial (for details see Yang et al; Leng and Ludwig (26) this issue). With these reservations in mind, we further suggest that the common denominator behind the contextual effects of OT is its role in regulating attention to social cues. Considering the role of the dopaminergic system in assigning salience and regulating attention, we argue that through its interaction with the dopaminergic system, OT has a major role in attention orienting to social cues. We discuss the implications of this model for the therapeutic applications of OT in conditions characterized with aberrant social behavior.

**Oxytocin, prosocial behavior and social salience**

Prosocial behaviors are broadly defined as acts that are viewed by one’s social group as beneficial to other people (27), and have been linked to social behaviors such as sympathy, empathy and cooperation (28). Consistent with the Kosfeld and colleagues (9) findings suggesting that inOT increases cooperative trust, subsequent studies showed that individuals treated with OT continued to exhibit trusting behaviors even after being betrayed (29), were more generous in a money gifting generosity game (10), and perceived others in ways that facilitate affiliation (12, 13). Furthermore, findings suggest a considerable role for OT in social cognition or component processes that contribute to social cognition. For example, inOT improved mentalizing (i.e., the ability to infer the mental states of others) (30), recognition of emotions from other people’s facial expressions (31, 32), as well as eye contact between
participants (33, 34). The positive effects of inOT on socio-cognitive abilities and functioning led researchers to evaluate its therapeutic applications in various psychiatric conditions such as autism, schizophrenia and personality disorder, yielding mixed results (35-40).

Although many of the findings support the prosocial effect of OT, a closer look at other results reveals the variable nature of its effects in humans. First, several studies report negative (i.e., non-prosocial) effects such as increasing feelings of mistrust (13, 41). Second, many of the OT’s previously reported effects appear to be moderated by situational or individual factors. For example, the prosocial effects of OT on trusting behavior (9) disappear if the other is portrayed as untrustworthy (42) or is unknown (13). Moreover, it has been shown that OT differentially affect male and female participants on social perception task (43), as well as whether the target was male or female (44).

Given that the non-prosocial effects of OT have been found in negative contexts involving threat (13) or competition (14) it may be argued that OT enhances prosocial behaviors only when the social context involves cooperative and positive emotions. Yet, in competitive, aggressive contexts, OT may enhance competitive or aggressive behaviors. Under such conditions, OT may increase the salience of threat signals, which may result in attention orienting responses to threatening social cues rather than positive social cues.

**Oxytocin, stress and social salience**

Several studies propose that the social effects of OT can actually be linked to its anxiolytic effects, whereby consequential stress reduction promotes prosocial behavior [for reviews see (45, 46)]. Indeed it has been repeatedly reported, both in rodents and humans, that the perception of threatening stimuli increases OT release and that the activation of the OT system regulates
behavioural and physiological manifestations of anxiety (for details see Neumann & Slattery (47) this issue). For example, Kumsta and Heinrichs (46) list in their review studies showing that *in OT* attenuates neuroendocrine stress reactivity (48), decreases amygdala activation in response to threatening stimuli (49), as well as an association between the OT receptor gene rs53576 and dispositional empathy and physiological stress reactivity (50). Furthermore, it has been shown that OT reduces anxiety in rats selectively bred for high versus low anxiety-related behaviour (51).

Recent findings, however, show that OT may actually facilitate stress (52), as well as protective behaviors including aggression regardless of stress reduction (53), thus indicating that OT may modulate social behavior in a manner that is not related to stress reduction. In this regard, it has been shown that the administration of OT increases territorial and maternal aggression in rats (54), and fails to reduce aggression in high anxiety rats (55). Similarly, there is evidence showing that OT is associated with increased stress and aggression in humans. For example, *in OT* increased the probability of aggression among high trait aggression people towards an intimate partner, as self-disclosed by participants, following a provocation task (16), and that OT increases anxiety to unpredictable threat (56). Moreover, contrary to what would be predicted by the OT stress theory, it has been reported that offenders who experienced early childhood maltreatment show raised urinary OT levels (a putative measure of OT) (57, 58) (However, see Heim et al. for contradictory findings (59)). Similarly, plasma OT levels correlated with higher social anxiety symptoms in patients with Generalized Social Anxiety Disorder (60), as well as with interpersonal distress (17).

Collectively, the studies reported here may indicate that OT may increase the salience of safety signals in positive supportive contexts (e.g. 48), which may attenuate stress. Conversely,
in unpredictable threatening situations (e.g. 56), OT may trigger orienting responses to threat rather than safe signals and increase anxiety.

**Oxytocin, inter-group relations and social salience**

A different line of studies suggest that the inconsistent findings in the literature can be explained by the preferences and predispositions of individuals toward ingroup members. Support for this account comes from studies demonstrating a role for OT in regulating parochial altruism (12), ethnocentrism and increase in-ingroup preference (61), in-group conformity (62), as well as empathic neural responses in the P2 time window (a fronto-central positive waveform component of the event-related potential activity at 128-188 ms) to pain expressions of racial ingroup members but not to racial outgroup members (63).

Interestingly, these studies have mainly reported that OT increases ‘in-group love’ rather than elevating ‘out-group hate’ (61, 64). This view is compatible with the social salience hypothesis presented here. Considering that humans have basic biases for in-group related information (65, 66), it is not surprising that OT preferably modulates emotions towards in-group members. Yet, when an out-group member is associated with a conflictual group rather than a neutral group, the salience of this member may be higher than an in-group member. In such contexts, OT would be expected to regulate emotions towards the conflictual out-group. This view is consistent with recent findings showing increased empathy for the pain of out-group members in the context of the Israeli-Palestinian conflict (35, 67).

Taken together, the above discussion of these various accounts demonstrates that the effect of OT is susceptible to contextual factors, individual differences and target. It also appears to have both positive/prosocial and negative/non-prosocial effects. In the next section, we first
provide evidence for the role of OT in modulating the salience of social cues and the variation of this modulation as a function of the individual’s characteristics. We then propose that OT exerts these effects by altering attentional neural mechanisms through its interaction with the dopaminergic system, which is involved in directing attention and assigning salience to relevant information.

The role of oxytocin in modulating salience of social cues

OT appears to exert its effect on social behavior in humans in a context-dependent manner. While it has been shown that OT increases trust, love and empathy to ingroup but not to outgroup members (12, 63), OT may also lead to defensive forms of aggression towards outgroup members (12) as well as significant others (16), suggesting that OT promotes aggressive tendencies toward the other depending on the nature of the relationship between the observer and the target. These findings are consistent with reports from animal research suggesting that the role of OT extend beyond simple approach functions, whereby it also regulates selective aggression against male intruders in pair-bonded prairie voles (68) and maternal aggression in postpartum and lactating rats (69).

The context-dependent effect of OT is also evident when processing socially human-specific information. Although it has been suggested that OT is not uniquely social and mainly modulates approach behaviors (70, 71), it has been reported that OT enhances memory for pictures of faces, but not for pictures of non-social objects (72), and affected arousal ratings to pictures of humans but not of animals (73). This social human-specific effect has gained yet another support from a study showing that OT induces a significant reduction in amygdala-
midbrain connectivity to fearful visual images, with a more prominent effect on socially salient stimuli (faces) as compared to nonsocial scenes (49).

The mechanism that may underlie the external environmental effects of OT may relate to modulation of attentional orienting responses. Indeed, it has been suggested that OT elevates both the number of saccades towards the eye region (33, 74), enhances stimulus-induced pupil dilation (22), increases orienting of attention in response to emotional gaze cues (75), as well as attentional shifts towards happy facial expressions presented for 100 ms (76). Taken together, these findings point to the possibility that OT modulates overt as well as covert attentional shifts at early stages of attentional processing, and suggest that the role of OT in modulating attentional orienting responses may underpin its role in increasing the salience of social cues.

Oxytocin’s modulation of salience as a function of the individual’s characteristics

The effects of OT are not uniform across all individuals and appear to vary as a function of gender, personality traits, attachment styles and psychopathology. For this reason, OT might affect individuals differently based on the specific constellation of traits and characteristics that form the person’s profile.

While much of the reported effects of OT are in male participants, a number of studies have shown that OT affects men and women differently. For example, in contrast to its effect in reducing amygdalar activation in men, inOT is associated with increased activation in women (20). Consistent with this finding, inOT increases the neural activation patterns of men to levels indistinguishable from the baseline activation patterns in women following placebo (77). It also increases the empathic responses of male participants to a level similar to that of untreated
women (78), and differentially affects social perception, such that it facilitates the accurate perception of competition in men, and kinship in women (43).

In addition, a large body of literature suggests a key role for OT in forming bonds and attachments between infants and caregivers. It has been shown that mothers with secure attachment have elevated OT levels compared to mothers with insecure/dismissing style in response to their infants at 7 months (79), and that mother-infant affect synchrony moderates the degree of the correlation between the mother and the infant salivary OT levels (80). In addition, studies have shown that inOT enhances the subjective experience of secure attachment in insecure male adults (81), and that individual differences in attachment was predictive of plasma OT levels in premenopausal women (82). Collectively, these findings suggest that individual differences in attachment style have consequences on the development of the oxytocinergic system and how it reacts to the administration of exogenous OT or to social behaviors that are typically associated with an increase in OT levels.

Moreover, inOT has been shown to selectively improve empathic accuracy in the less socially capable individuals (19) as well as mentalizing abilities in individuals with impaired empathy (83), suggesting that OT may have a more limited role in augmenting social salience, one that perhaps preferentially benefits those with lower baseline capabilities. Along the same line, Clark-Elford et al (84) have recently shown that while OT reduces the difference between individuals with high and low social anxiety in attentional bias for emotional faces, this effect was driven by increasing the bias in the controls. The absence of such effects in individuals with high social sensitivity suggests that OT may have a saturation point in that it may not make available information that previously (i.e., under the placebo condition) is already salient (however see (85)).
Oxytocin, dopamine and social salience

Salience is a key attentional mechanism associated with the ability to reorient to (or filter out) salient stimuli. The detection of salient stimuli is centrally regulated by the dopaminergic system (86, 87) and is characterized by a sharp increase (phasic) in dopaminergic activity, 70-100ms post the presentation of the stimulus, and occurs before an orienting gaze-shift, which has a latency of 150-200ms indicative of its role in attention reorienting (88). This signal is involved in rapid detection of and alerting to potentially important sensory cues (89), and is most likely triggered by information received from the superior colliculi (90)—brain stem areas associated with eye movement and attention. Importantly, the sensitivity of DA neurons to this information is dependent on basal levels of tonic DA and is modulated by homeostatic biological functions as well as individual characteristics (91). Specifically, a reduction in tonic DA release can lead to homeostatic compensations that would ultimately result in exaggerated phasic dopamine release to salient stimuli. These alerting signals are sent to salience-coding DA neurons in the VTA, and from the VTA to mesolimbic structures including the amygdala (most likely the central nucleus of amygdala), the nucleus accumbens, which are equally responsive to aversive and rewarding stimuli, to assess their value and valence (89). It is therefore possible that changes in phasic DA may account for the external contextual effects of OT. Furthermore, it is possible that given that tonic DA activity contributes to individual differences and psychopathology (92), the effect of OT depends on the person’s individual characteristics.

Indeed, accumulating evidence suggests that these dopaminergic-mediated effects are modulated by the oxytocinergic system. Animal studies have shown that OT has numerous binding sites across the mesocorticolimbic DA system (93-96). In addition, a sizeable animal
literature has demonstrated that both DA and OT interactively affect various social behaviors in key regions within this system [for reviews see (6, 97-99)]. For example, the seminal work of Liu and Wang has shown that coactivation of D$_2$ receptors and OT receptors in the nucleus accumbens is necessary for pair bond formation and maintenance (100), which recently has been suggested to be facilitated by the presence of DA D$_2$ receptor-OT receptor (D$_2$R-OTR) heteromers (101). In addition, OT has been shown to have a direct effect on the release of DA within the mesocorticolimbic system, including the VTA and the nucleus accumbens in female rats during pup grooming (102), suggesting that OT enhances the salience of social cues by boosting dopamine’s coding signal.

In humans, OT has a myriad of binding sites across the brain, including limbic and autonomic areas (103, 104), and the interaction between the oxytocinergic and dopaminergic systems has been reported in several studies examining the association of OT with activity within dopaminergic regions. For example, research has shown that OT enhances responses in brain reward regions including the nucleus accumbens and VTA in pair-bonded males who were exposed to pictures of their female partner when compared to unfamiliar women (105), as well as within the caudate nucleus (another reward structure) and the amygdala in response to reciprocated cooperation (106). However, OT has been shown to also increase activity in the VTA in response to both positive/rewarding (friendly faces) and negative/punishing (angry faces) social cues (107), suggesting that OT attaches salience to social cues irrespective of their valence. More recently, a positron emission tomography study suggests that OT enhances attractiveness to unfamiliar faces independent of the DA reward system (108). However, this study also observed increased D2 DA binding in the dorsomedial prefrontal cortex (part of the
mesocorticolimbic system), which has been shown to be involved in attending to stimuli independent of valence or type of stimuli (109).

Based on the above discussion, the interaction of OT and DA within the mesocorticolimbic system, which is centrally involved in the processing of aversive and rewarding events, the assignment of salience and attention reorienting, suggests that OT has a key role in these functions [see (110) for a similar view]. Specifically, we propose that OT modulates the salience of social stimuli by regulating dopamine’s salience-coding and attention reorienting signal (See Figure 1). The amygdala (and particularly its central nucleus) is perhaps the most likely site where the interactive effect of DA and OT on salience and attention reorienting takes place, given its established role in attention reorienting and the assignment of salience to social and positive and negative emotional stimuli (111-115). Evidence for this interaction in humans has recently been reported in an fMRI study showing that amygdala activation in response to social stimuli was interactively modulated by the CD38 gene (which is involved in OT secretion) and the catechol-o-methyltransferase (COMT) genotype (which is involved in the degradation of DA) (116). Intriguingly, there is some evidence that CD38 gene polymorphisms in humans can explain individual variability in the accuracy of detection of another’s direction of gaze(117). In addition, research has shown that in OT increases gaze shifts to the eye region (irrespective of the emotional expression) and that this gazing pattern is associated with enhanced functional coupling of the posterior amygdala and the superior colliculi (74).

**Figure 1 about here**

By framing the effects of OT in terms of its modulatory role of attention reorienting and the assignment of salience through its interaction with the dopaminergic system, we highlight
that the behavioral effects of OT are highly dependent on the degree to which social cues are made irrelevant or salient and relevant. However, further research is required to clarify the role of OT in enhancing social salience by using attention-reorienting/divided attention paradigms that directly test its effects on the processing of relevant social cues in the presence of irrelevant but salient social and non-social stimuli. In addition, the role of the amygdala in facilitating the interactive effect between the dopaminergic and oxytocinergic systems combined with its massive network of afferent and efferent projections to cortical and subcortical regions underscores its central role in mediating negative and positive social effects. This is consistent with evidence suggesting that a disruption to the oxytocinergic-dopaminergic system prevents the integrated signaling between the amygdala and the prefrontal cortex, which consequently leads to impaired attention reorienting, social perception, and socio-cognitive abilities (112).

Oxytocin and other neurotransmitters

Our model has focused on the interaction of OT with the DAergic system. However, OT also interacts with other neurotransmitters including acetylcholine, glutamate, GABA, and serotonin (118, 119). For example, receptor mapping studies in nonhuman primates have shown that OT receptors are concentrated in cholinergic rich brain areas that modulate visual attention (nucleus basalis of Meynert and superficial grey layer of superior colliculus) as well as areas involved in auditory processing (e.g., trapezoid body) (120, 121). In addition, animal research documents that social interaction in mice require the coordinated activity of OT and serotonin in the nucleus accumbens (122), as well as in humans whereby OT appears to exert inhibitory effects in the
dorsal raphe nucleus, the core area of serotonin synthesis, and in regions within the mesocorticoclimbic systems which include the amygdala/hippocampal complex, the insula and the orbitofrontal cortex (123). While these interactions challenge our ability to identifying which of these transmitters are most relevant for OT-mediated functions, they highlight the complex neuro-machinery harnessed by the OT system to exert its effects. Therefore, uncovering the cascading effects of these transmitters would be key to understanding their role in mediating oxytocinergic functioning.

**Implications for psychopathology**

The therapeutic potential of OT [for a recent review see (124)] must be considered in light of its potential negative effects. Within the context of our model, administering OT, for example, to individuals exposed to aggressive or threatening contexts would be expected to increase the salience of threat signals. This may, in turn, trigger high levels of anxiety, which may be problematic for individuals with anxiety disorders. Similarly, treating an individual with OT in uncontrolled situations could enhance the salience of negative social interactions and thus chronic, daily OT therapy without the assurance of a positive social experience may not be the best course of action.

The therapeutic potential of OT, has most prominently been discussed in relation to autism and schizophrenia. Rosenfeld and colleagues (112) have proposed that a breakdown in the oxytocinergic-dopaminergic-amygdalar system may underlie social salience processing and attentional dysfunctions in autism and schizophrenia spectrum disorders. However, these conditions appear to respond differently to salient social cues (125). More specifically, individuals with autism show deficits in attending to salient social stimuli (126), whereas
individuals with schizophrenia show difficulty in suppressing salient information (127). Thus, we suggest that understanding how various conditions affect the processing of socially salient information is key to assessing the therapeutic potential of OT.

Concluding remarks

OT plays a major role in social behavior. Its effects, however, appear to be context dependent, and can be both positive and negative. In this review, we have proposed a theoretical framework to account for the variegated role of OT in human behavior by invoking its strong functional interaction with the dopaminergic system and thus its role in increasing the salience of social cues. This account may offer a more nuanced framework to assessing the potential therapeutic applications of OT in different populations and settings.

Financial Disclosures

Dr. Shamay-Tsoory reported no biomedical financial interests or potential conflicts of interest.

Dr. Abu-Akel reported no biomedical financial interests or potential conflicts of interest.

References

42. Mikolajczak M, Pinon N, Lane A, de Timary P, & Luminet O (2010) Oxytocin not only increases trust when money is at stake, but also when confidential information is in the balance. *Biological psychology* 85(1):182-184.


Figure Legend

**Figure 1.** Simplified schematics of the interaction between the dopaminergic and oxytocinergic systems in the presence of an external visual cue. Blue circuit represents attention-orienting/assignment of salience and the red circuit represents the oxytocinergic system and its modulatory role of the dopaminergic system. The responsivity of the dopaminergic system and thus its modulation by the oxytocinergic system is dependent on the availability of tonic dopamine (not shown) whose levels are determined by homeostatic biological functions and individual characteristics such as sex and degree of psychopathology. The appraisal of salient stimuli takes place both in the amygdala and the nucleus accumbens via an interactive dopaminergic-oxytocinergic mechanism. This information is relayed to the VTA to regulate DA release, which is conjointly modulated by direct projections from the oxytocinergic system, as well as to prefrontal regions for higher cognitive processing. Unlike the nucleus accumbens, the amygdala is additionally involved in communicating attentional responses to the superior colliculus to shift or maintain focus of attention. The projection from the frontal eye field (FEF) to the superior colliculus represents one route through which top down control of attention reorienting might take place. SC= Superior Colliculus; VTA= Ventral Tegmental Area; NA=Nucleus Accumbens; PFC= Prefrontal Cortex; FEF= Frontal Eye Field.