

# Opposing Effects of Oxytocin on Overt Compliance and Lasting Changes to Memory

Micah G Edelson<sup>\*1,6</sup>, Maya Shemesh<sup>1,6</sup>, Abraham Weizman<sup>2,3</sup>, Shahak Yariv<sup>1,4</sup>, Tali Sharot<sup>5</sup> and Yadin Dudai<sup>1</sup>

<sup>1</sup>Department of Neurobiology, the Weizmann Institute of Science, Rehovot, Israel; <sup>2</sup>Research Unit, Geha Mental Health Center and Felsenstein Medical Research Center, Beilinson Campus, Petach Tikva, Israel; <sup>3</sup>Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; <sup>4</sup>Department of Psychiatry, Haemek General Hospital, Afula, Israel; <sup>5</sup>Department of Cognitive Perceptual and Brain Science, University College London, London, UK

From infancy we learn to comply with societal norms. However, overt compliance is not necessarily accompanied by a change in internal beliefs. The neuromodulatory processes underlying these different phenomena are not yet understood. Here, we test the role of oxytocin in controlling overt compliance versus internalization of information delivered by a social source. After intranasal oxytocin administration, participants showed enhanced compliance to the erroneous opinion of others. However, this expression was coupled with a decrease in the influence of others on long-term memories. Our data suggest that this dissociation may result from reduced conflict in the face of social pressure, which increases immediate conforming behavior, but reduces processing required for deep encoding. These findings reveal a neurobiological control system that oppositely affects internalization and overt compliance. *Neuropsychopharmacology* (2015) **40**, 966–973; doi:10.1038/npp.2014.273; published online 5 November 2014

## INTRODUCTION

When faced with social pressure, it is often adaptive to conform, to avoid rejection, ridicule and punishment (Asch, 1952; Kelman, 1961; Smith and Mackie, 2007). However, when conforming, the individual can act in two different ways. Often it is advantageous to internalize information provided by others, as they can be excellent sources of information (Boyd *et al*, 1988; Haun *et al*, 2013). However, at times the majority may be wrong. Furthermore, frequent alterations to beliefs in response to external pressures may lead to maladaptive changes in the individual's self-image or a damaged sense of personal identity (Kelman, 1961; Rothbaum *et al*, 1982). Internalization is also not an optimal strategy under circumstances involving a high degree of environmental uncertainty and unstable social structures (Boyd *et al*, 1988; Haun *et al*, 2013). Thus, an individual must at times comply with societal constraints without internalization (Asch, 1952; Haun *et al*, 2013; Smith and Mackie, 2007). This type of behavior, for example, is prevalent in newcomers joining existing social groups (Moreland and Levine, 1989). There is evidence that such processes can malfunction in several psychopathologies (eg, (Marsella, 1975)). For example, it has been proposed

that autistic children do not properly internalize shared experiences (Rhode, 2012).

Internalization in response to social pressures (also called Private conformity (Smith and Mackie, 2007)) can be measured by exploring memory modification following the application of social influence and its subsequent removal (Edelson *et al*, 2011, 2014; Meade and Roediger, 2002; Smith and Mackie, 2007). In this case, even when social influence is removed, the individuals will persist in demonstrating support for the option endorsed by the group, as the memory has already internalized and is perceived as one's own (Kelman, 1961; Meade and Roediger, 2002; Smith and Mackie, 2007). Overt compliance (also known as Public conformity), by contrast, does not involve internalization and is dispelled when the veracity of the socially transferred information abates (Asch, 1952; Kelman, 1961; Smith and Mackie, 2007). As conformity can have a significant societal impact, for example, in regard to totalitarian groups and regimes (Kelman, 1961; Smith and Mackie, 2007), understanding the underlying neurobiology of these two distinct processes is important.

The neuromodulator oxytocin has been implicated in mediating a wide variety of social behaviors, including the facilitation of social bonds, trust, and affiliative behavior (Bartz *et al*, 2011; Carter, 2013; Feifel, 2012; Feldman, 2012; Guastella *et al*, 2013; Harris and Carter, 2013; Heinrichs and Domes, 2008; Hollander *et al*, 2007; Meyer-Lindenberg *et al*, 2011; Neumann, 2008; Shalvi and De Dreu, 2014; Stallen *et al*, 2012; Young and Wang, 2004). Moreover, oxytocin has been found to affect both immediate and long-term reactions to information from a social source (Bartz *et al*, 2011; Bruins *et al*, 1992; Carter, 2013; Feldman, 2012;

\*Correspondence: Dr MG Edelson, Neurobiology, Weizmann Institute of Science, Neurobiology department Weizmann institute, 234 Herzl St, Rehovot, Select state/region 7610001, Israel, Tel: +972 8 934 3711, Fax: +972 8 946 9244, E-mail: micahedelson@gmail.com

<sup>6</sup>These authors contributed equally to this work.

Received 15 July 2014; revised 28 September 2014; accepted 1 October 2014; accepted article preview online 13 October 2014

Geenen *et al*, 1988; Guastella *et al*, 2013; Harris and Carter, 2013; Neumann, 2008; Rimmele *et al*, 2009). Recent animal studies suggest that oxytocin may even exercise opposite influences in these two circumstances (Bales *et al*, 2013; Huang *et al*, 2013). These findings make oxytocin a prime candidate for influencing internalization, which will result in persistent long-term effects (Kelman, 1961; Smith and Mackie, 2007), and/or transient responses to social influences. Here, we test whether these processes can be controlled via oxytocinergic modulation.

## MATERIALS AND METHODS

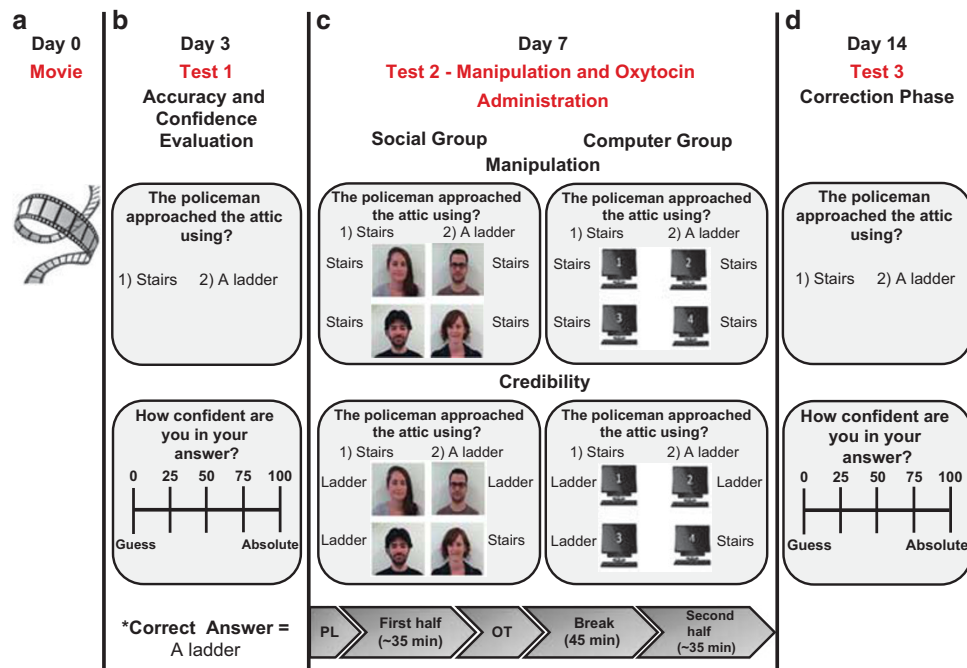
### General Outline

In a four-phase protocol, 92 male participants were exposed to erroneous information, while receiving placebo or oxytocin in a within-subject randomized controlled crossover design (Figure 1). Participants first viewed an eyewitness styled documentary movie, in groups of 6–8 people (Figure 1a). For increased efficiency in ‘minimal group’ settings, where groups are arbitrarily constructed based on trivial criteria, we enhanced in-group affiliation (see (Stallen *et al*, 2012)), by instructing participants to introduce themselves (name, age and occupation) and by informing them that they were part of the same group (‘Blue group’), whose performance would be compared with other out-groups. All subsequent experimental stages were performed

by participants individually in the lab. Three days after viewing the film, the participants completed a memory test (Memory Test 1; 400 two-forced choice questions), which served to assess their baseline memory accuracy and confidence before the manipulation stage (Figure 1b).

Four days later, the participants answered the same memory questions (Test 2; Manipulation phase; Figure 1c). On this occasion a manipulation was introduced in an attempt to induce conformity. Before answering each question in this second test, participants were presented with answers supposedly given by four of their fellow group members (46 participants; Social group), or by four computer algorithms (46 participants; Computer group). This latter group served to assess the social specificity of the pharmacological manipulation. In a subset of questions, originally answered correctly by the target participant with medium confidence (as identified by Test 1), the four external answers were all false (Manipulation questions, 60–80 questions). To maintain a credible manipulation, we added Credibility questions in which the patterns of external answers supported the participant’s initial answer and confidence (see Supplementary Information for more details). Placebo and oxytocin were administered intranasally at this Manipulation stage in an intra-subject crossover design (24 International Units; Figure 1c; see Substance Administration below).

One week following the manipulation stage, the participants returned to the lab and were informed that the four



**Figure 1** Experimental protocol. (a) Participants watched a movie in groups of 6–8. (b) Three days later, participants answered memory Test 1 to assess baseline accuracy and confidence. Questions answered correctly with medium confidence were chosen as Manipulation questions in Test 2. (c) Four days later, in memory Test 2, the Manipulation was introduced and oxytocin was administered. Before being able to answer, the participants viewed four fabricated answers of their group members (Social group) or four computer algorithms (Computer group). In Manipulation questions all four answers were incorrect. The rest of the questions insured the plausibility of the manipulation and included the Credibility questions, where different answer patterns were presented supporting the participant’s view. *Bottom panel*: Oxytocin administration in a crossover within subject design. Placebo was administered before the first half of the test. Oxytocin was administered before the second half of the test. The order of administration was not a sufficient explanation for the results because this is controlled for by the Computer group and the No-information control condition (for further information see Supplementary Information). (d) Memory Test 3 served to test Persistent errors by eliminating the immediate overt motivation to conform.

external answers given in the previous session were in fact determined randomly (ie, equal probability of a correct or incorrect answer). This rendered the information previously provided equivalent to an uninformative guess, and not to a deliberately misleading response. Discrediting the source of information in this manner is an efficient way to distinguish Private and Public conformity (Allen, 1975; Asch, 1952; Edelson *et al*, 2011, 2014; Haun *et al*, 2013; Smith and Mackie, 2007). This is because such a procedure eliminates both Informational and Normative motivations for immediate compliance (Deutsch and Gerard, 1955) and should thus assess only the internal representation of the individual (Allen, 1975; Edelson *et al*, 2011; Haun *et al*, 2013; Kelman, 1961; Smith and Mackie, 2007).

In addition, this procedure emulates realistic situations in which the credibility of information is refuted (eg, testimony and evidence provided to jurors, which is later classified as incorrect or inadmissible (Lewandowsky *et al*, 2012)) and natural changes prevalent in the dynamic human social environment (Byrne and Whiten, 1989). Participants understood that previous information provided was random and not biased for or against the real participant/computer answers (for additional controls and discussion see Supplementary Information). The participants were then requested to complete the memory test again based on their original memory of the movie (Test 3; Correction phase; Figure 1d). Finally, participants filled out trait questionnaires and were debriefed (see Supplementary Information).

Events of interest (Manipulation questions) were divided into three conditions according to the participant's answers on the three memory tests. (a) Non-conformity: events in which the participant maintained a correct answer in the manipulation phase (Test 2), despite the contradictory external information. (b) Transient errors (Public conformity): events in which the participant outwardly complied in the manipulation stage by endorsing the wrong answer (Test 2), but reverted back to the correct response once the influence was removed (Test 3). (c) Persistent errors (Private conformity): events in which the participant's conformity to misleading influence in the manipulation stage (Test 2) persisted even after the influence was removed (Test 3), reflecting internalized long-lasting alteration to beliefs.

## Participants

In accordance with the majority of previous literature (MacDonald *et al*, 2011; Meyer-Lindenberg *et al*, 2011), only male participants were recruited. This is due to potential side effects on females including ethical and legal repercussions, and to avoid confounds related to gender and the menstruation cycle ((Carter, 2013; MacDonald *et al*, 2011; Owen and Hauth, 1996) for more details see Supplementary Information). Each participant was randomly assigned to either the Social or Computer group ( $N=92$ , mean age  $27.1 \pm 0.7$  (SEM)).

In addition, in accordance with previous literature (Ortmann and Hertwig, 2002; Stang, 1976), only participants who did not indicate suspicion of the manipulation during debriefing were included in the final analysis. This is because the effects of external influence cannot be accurately examined if the subjects do not believe the

answers are real (Ortmann and Hertwig, 2002; Stang, 1976). Exclusion criteria (detailed in Supplementary Information) resulted in a final 31 participants included in the Social group (mean age  $26.8 \pm 0.9$ ) and 32 participants included in the Computer group (mean age  $27.6 \pm 0.9$ ). This exclusion rate was well within the norm of previous equivalent studies (for reviews see (Ortmann and Hertwig, 2002; Stang, 1976)). Critically, adding the excluded participants to the analyses reported in result section did not significantly alter the results (see Supplementary Information). Thus, the exclusion criteria cannot account for the reported results.

All participants signed informed consent forms and were paid for their participation. The study was in accordance with the provisions of the World Medical Association Declaration of Helsinki and was approved by the Institutional Review Board of the Geva Mental Health Center, Petah Tikva, Israel, and limited to 100 participants, thus determining the sample size (ie, adding an additional minimal group of participants (six participants for each group) would total over 100).

## Substance Administration

We followed the recommended guidelines for the standardization of oxytocin administration (Guastella *et al*, 2013; MacDonald *et al*, 2011). Participants in both the Social and Computer groups self-administered an intranasal dose of 24 international units (IU) oxytocin (Syntocinon Spray, Novartis, Basel, Switzerland; 3 puffs per nostril, alternately, with one minute interval, each with 4IU oxytocin) or placebo, under the experimenter supervision. The placebo vehicle solution was manufactured by 'Maayan-Haim' pharmacy (Bet Dagan, Israel) and contained the same amounts of all ingredients present in the oxytocin solution (except for the neuropeptide). Subjects were not aware that they received both placebo and oxytocin, and, when debriefed, subjects could not identify which administration contained the treatment over chance levels, nor did they report differences in perception or physical sensation between the treatment and the placebo (see Supplementary Information). To maximize effectiveness of oxytocin administration, participants were given a 45-min break between the two experimental sessions, a sufficient time period in order for the drug to take effect (van IJzendoorn and Bakermans-Kranenburg (2012); MacDonald *et al*, 2011). For additional controls related to the order and context of the pharmacological administration see Supplementary Information.

## Statistical Analysis

Data were analyzed using R (The R Project for Statistical Computing), MATLAB software (2012b, MathWorks Inc., Natick, MA) and SPSS (IBM Corp. 2012, Version 21.0. Armonk, NY). Effects are reported as means. Unbiased effect sizes for *t*-tests (all *t*-test are within subject) are reported after applying Hedges correction to Cohen within-subject  $d'_{rm}$  (Hedges  $g_{rm}$ ), using the spreadsheet provided by Lakens, 2013. Standardizer for Cohen  $d'_{rm}$  effect size was selected according to appropriate standard deviation of the difference scores (Cumming, 2013; Lakens, 2013). Effect sizes for ANOVA are reported as partial Eta squared ( $\eta^2_p$ ).

The size of the effect is interpreted according to recommendation in literature (ie, when effect size cannot be compared with equivalent study in literature, rule of thumb is 0.2, 0.5 and 0.8 for small, medium and large effects, respectively, for  $g_{rm}$ , and 0.01, 0.06 and 0.14 for small, medium and large effects respectively for  $\eta^2_p$  values; (Cohen, 1988; Lakens, 2013)). All confidence intervals (CI) were calculated via a bootstrapping approach (Efron and Tibshirani, 1993) with 100 000 simulations, which does not assume any specific a-priori distribution. As an additional conservative measure, we account for selective inferences related to the replicability of the results, by using a false coverage rate (FCR) approach to correct our 95% CIs for findings selected from various variables tested (Benjamini *et al*, 2009) and the reported CIs include this correction (FCR CI). This excludes the mediation model that was constructed post-hoc according to observed data using R programming software and PROSESS SPSS macro, model 4 (Hayes, 2013) with 10 000 simulations and 90% CIs. CIs on repeated measures are on the appropriate difference of the means (Cumming, 2013).

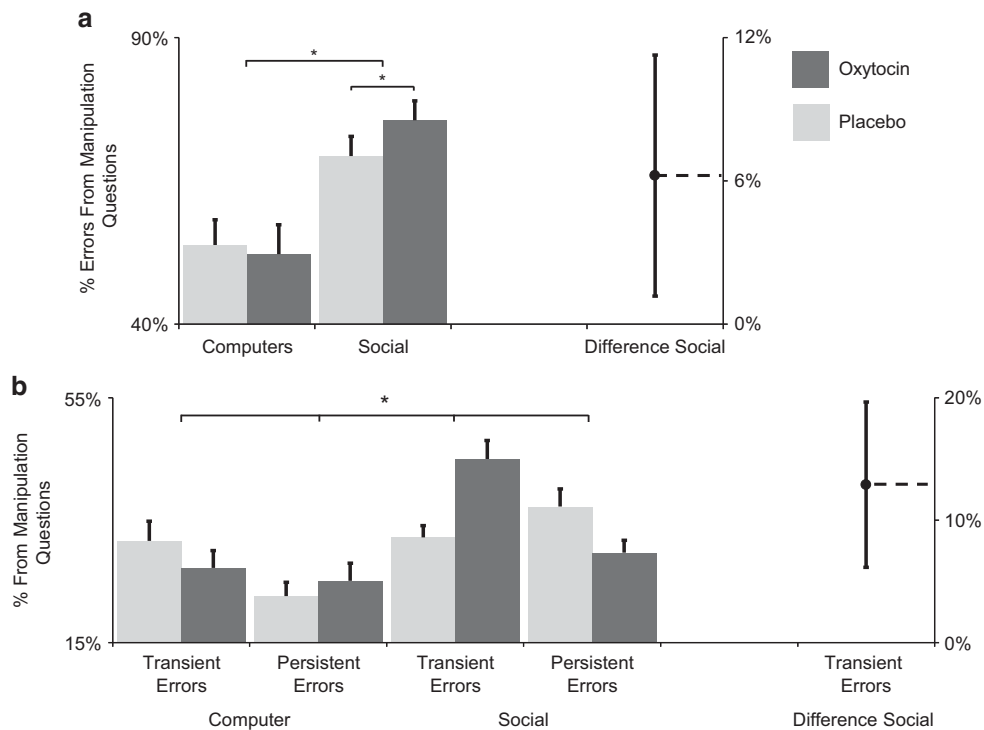
To increase interpretability of the results, we additionally report the common language (CL) effect indicator (Lakens, 2013) that expresses the probability that a randomly sampled person has a higher value on one measurement than the other in the current study. We also ran 100 000 bootstrap simulations for half our population size (ie, randomly drawing  $N = 16$ , per measure tested, recording the

mean of the sample, and repeating this procedure 100 000 times to create a distribution of means), we then calculate the percentage of these partial sample means (PPM), falling below the confidence interval of our main reported effect. This represents a measure of the internal reliability of our main effect given half the sample size.

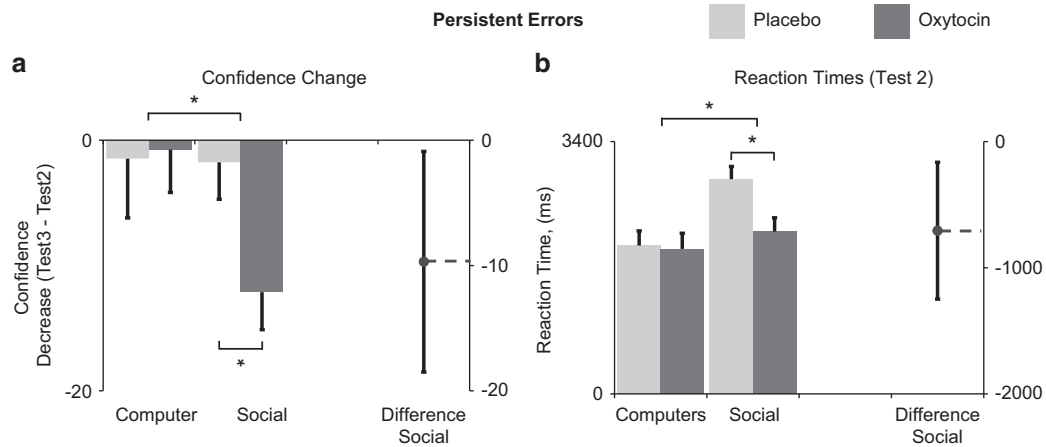
## RESULTS

### Oxytocin Effects on Immediate Compliance

Our data (Figure 2a) indicated that oxytocin led to an increase in conforming behavior (adopting the erroneous response endorsed by the group in Test 2) *vs* ignoring the social information (Non-conformity). This was evident only under conditions of social influence but not when misinformation was conveyed by an inanimate source (interaction of group (Social/Computers) and substance (oxytocin/placebo);  $F_{(1,61)} = 4.2$ ,  $p < 0.05$ ,  $\eta^2_p = 0.07$  driven by an increase in conformity under oxytocin *vs* placebo in the Social group; mean increase 6.2%, 95% FCR CI [1.2, 11.3], *post hoc t-test*,  $t(30) = 2.8$ ,  $p < 0.05$ ,  $g_{rm}$  effect size = 0.3, PPM = 5.4%, CL = 69%). For inter-individual differences in this measure see Supplementary Information and Supplementary Figure S1. This result is in line with a recent report demonstrating that oxytocin increases immediate social conformity (Stallen *et al*, 2012). However, a crucial question is whether the observed enhancement



**Figure 2** Oxytocin effects on conformity. (a) Conformity levels during social influence (Test 2, depicted in bar chart), demonstrated an interaction between group (Social/Computer) and substance (oxytocin/placebo) induced by an increase in conforming behavior under oxytocin in the Social group. Non-conformity levels are the inverse of conformity levels (ie, 1-% conformity). (b) Decomposition of this effect shows an interaction between group (Social/Computer) and substance (oxytocin/placebo) in the relative percentage (ratio) of transient compliance *vs* persistent errors, driven by an increase in transient errors in the Social group accompanied by a decrease in persistent errors (depicted in bar chart). \* $p < 0.05$ . Bar chart depicted with SEM. Blue dot and accompanying 95% false coverage rate (FCR) confidence interval correspond to the differential value as depicted on right axis. Confidence intervals are extended to include FCR correction for multiple comparisons.



**Figure 3** Oxytocin effect on persistent errors independently of transient errors. (a) Oxytocin reduced the robustness of persistent errors as indicated by a larger confidence decrease after social influence was removed in Test 3 (bar chart). (b) Oxytocin reduced reaction times during exposure to social influence in Test 2 (bar chart). In both cases, the interaction is driven by a significant effect in the Social group. \* $p < 0.05$ . Bar chart depicted with SEM. Blue dot and accompanying 95% false coverage rate (FCR) confidence interval correspond to the differential value as depicted on right axis. Confidence intervals are extended to include FCR correction for multiple comparisons.

in external tractability represents merely overt compliance (ie, transient errors) or an ingrained change in beliefs (ie, persistent errors).

As the number of conformity events is finite, the numbers of persistent and transient errors in each condition are partially dependent measures. Therefore, to test whether oxytocin induced a change in the relative occurrence of transient vs persistent errors, we calculated the ratio of transient vs persistent errors from all manipulation trials. We found that oxytocin modulated the proportion of immediate outward compliance events relative to long-term *bona fide* change (Figure 2b, interaction of group (Social/Computers) and substance (oxytocin/placebo);  $F_{(1,61)} = 5.9$ ,  $p < 0.05$ ,  $\eta^2_p = 0.09$ ). These results indicate that oxytocin altered the nature of conformity toward more transient manifestations, increasing transient errors in the Social group (Figure 2b, 12.9% vs the placebo baseline, FCR CI [6.7, 20.2],  $t(30) = 4.3$ ,  $p < 0.01$ ,  $g_{rm} = 0.94$ , PPM = 6.6%, CL = 78%), and decreasing persistent errors (Figure 2b, -7.5% vs the placebo baseline, FCR CI [-0.3, -14.9],  $t(30) = 2.1$ ,  $p < 0.05$ ,  $g_{rm} = 0.53$ , PPM = 6.7%, CL = 65%). Thus, the increase in general conformity levels under oxytocin in the Social group (Figure 2a) cannot be explained by an increase in persistent errors and must be due to an oxytocin-related enhancement of transient compliance (beyond oxytocin's effect on persistent errors). However, whether there is an accompanied decrease in the strength of persistent errors requires testing with behavioral measures independent from transient errors, which we report next.

### Oxytocin Affects Internal Representations Independently from Transient Errors

We previously demonstrated that the strength and resilience of a persistent error can be represented by the degree to which confidence in this wrong answer changes after the misleading influence is lifted (Test 3, Correction phase, (Edelson et al, 2011, 2014)). Simply put, when the loss of

external support leads to a large confidence decrease, this indicates that the erroneous internal representation is less robust. We now find that oxytocin administration affected the size of the confidence change after the removal of the manipulation (Figure 3a, defined as confidence Test 3 – confidence Test 2;  $F_{(3,73)} = 2.9$ ,  $p < 0.05$ ,  $\eta^2_p = 0.11$ ). This result was driven by a larger confidence decrease related to oxytocin administration in the Social group after social influence was lifted (mean relative to placebo = -9.7, FCR CI [-1.1, -18.7],  $t(25) = 2.5$ ,  $p < 0.05$ ,  $g_{rm} = 0.68$ , PPM = 3.0%, CL = 69%). For inter-individual differences in this measure see Supplementary Information and Supplementary Figure S1. Note that this result, which is independent from transient errors, cannot be explained by the initial level of confidence in Test 2, which was included as a covariate. Moreover, confidence in Test 1 as well as confidence in Test 2 did not show an interaction between conditions and groups (interaction  $F_{(1,61)} = 0.4$ ,  $p = 0.53$ , and  $F_{(1,61)} = 0.45$ ,  $p = 0.58$ , respectively,  $\eta^2_p < 0.01$ ), whereas confidence ratings in Test 3 showed an interaction ( $F_{(1,61)} = 6.2$ ,  $p < 0.05$ ,  $\eta^2_p = 0.09$ ), indicating the confidence differences in either of the two first tests are not a sufficient explanation for the results.

Our data show that oxytocin reduces the strength of persistent social influence. What is the etiology of such a decrease? One possibility is that oxytocin leads individuals to feel closer to the opinion of the group. This would increase immediate compliance and would diminish the conflict and mental deliberation associated with the dissonance between the individual's belief and the group's opinion (Izuma, 2013). However, such attenuated conflict may reduce deliberate attention processes and mnemonic operations required for robust long-term memory, resulting in weaker encoding and consolidation of the erroneous social influence (Cowan, 1995; Dudai, 2002).

Following this rationale, participants may be quicker to endorse persistent errors under oxytocin relative to placebo. An ANOVA with reaction time (RT) on Test 2 (of persistent

errors; ie, trials where subjects endorsed the false answers of the group in Test 2 and 3) revealed an interaction of group (Social/Computers) and substance (oxytocin/placebo) ( $F_{(1,61)} = 4.5$ ,  $p < 0.05$ ,  $\eta^2_p = 0.07$ ). The effect was driven by a decrease in RTs under oxytocin vs placebo in the Social group (Test 2; RT mean difference =  $-706$  ms, FCR CI [ $-128$ ,  $-1212$ ],  $t(30) = 2.9$ ,  $p < 0.05$ ,  $g_{rm} = 0.7$ , PPM = 4.5%, CL = 70%, Figure 3b). In other words, oxytocin reduced the time necessary to make a decision when faced with social pressure (Manipulation phase), an effect unique to social influence rather than a generalized motor effect (see also Shalvi and De Dreu, 2014). Note that RTs were in general longer in the Social group (average over all events and conditions 3173 ms [3646, 2700] and 2602 ms [3161, 2044] for the Social and Computer groups, respectively), and thus the RTs under oxytocin in the Social group should not be directly compared with RTs in the Computer group. The described interaction, however, is not sensitive to this difference.

As a supplemental measure, we also created a mediation model (Hayes, 2013; MacKinnon *et al*, 2007) with RT mediating between the substance administration factor (oxytocin/placebo) and the percentage of persistent errors from conformity questions. We find that the indirect effect representing the mediation was significant (unstandardized Beta =  $-0.018$ , [ $-0.051$ ,  $-0.001$ ], Kappa-Squared measure for effect size = 0.06 [0.01, 0.16]; for more information on mediation models and parameters see (Hayes, 2013; Preacher and Kelley, 2011)). These results further suggest that reaction times (or a measure correlated with RTs, eg conflict) may mediate the effect of oxytocin on persistent errors. Other factors, however, may also contribute as the relationship between substance administration and persistent errors remained significant after the addition of RT to the regression model ( $t = 2.6$   $p = 0.01$ , (Hayes, 2013)).

The sizes of our main effects of interest ( $g_{rm} = 0.9$  for increased transient error occurrence and  $g_{rm} = 0.7$  for both the decrease in persistent error confidence and reaction times) should be considered as medium to large both when comparing to common practice in literature (see methods) and to the effect size found in the only comparable study ( $d_{rm} = 0.6$ , calculated from data published in Stallen *et al*, 2012). This is also evident in our confidence intervals, although they are conservatively large due to an applied FCR correction meant to address replicability issues (Benjamini *et al*, 2009). The common language (CL) indicator reported helps interpret the results. We found that the CL effects range is 69–78% leading us to predict that  $\sim 7$  out of 10 males in the population would show our reported effects, a substantial fraction considering the current clinical applications of intranasal oxytocin.

## DISCUSSION

Healthy individuals are highly influenced by their interpersonal environment and often conform to societal norms (Haun *et al*, 2013; Izuma and Adolphs, 2013; Smith and Mackie, 2007). However, overtly compliant behavior is not always accompanied by a change in the individual's internal beliefs (Haun *et al*, 2013; Kelman, 1961; Meade and Roediger, 2002; Smith and Mackie, 2007). For example,

candidates in a job interview may comply with the expected behavior, although it does not necessarily reflect their internal beliefs (Kelman, 1961). This dissociation is important because otherwise individuals would experience constant fluctuations in internal representations and a damaged sense of personal consistency (Baumeister, 1982; Smith and Mackie, 2007).

Our results are, to the best of our knowledge, the first description of a neuromodulatory-mediated process subserving a balance between transient compliance and internalization of social information. We found that oxytocin increased immediate compliant behavior, a finding consistent with the results of a previous report testing oxytocin effects on the initial stages of social influence (Stallen *et al*, 2012). However, we also found that oxytocin had the opposite effect on long-term memory representations, reducing the robustness of persistent errors. This was not the result of the order of administration, nor a general memory context effect, as it was evident only in the Social group and not when information was attributed to an inanimate source (for additional controls relating to order of administration see Supplementary Information). Furthermore, this was not simply a relaxation to baseline levels as the reduction in robustness of persistent errors under oxytocin was larger than under placebo.

As it is not plausible that externally administered oxytocin remains in the central nervous system until the Correction phase (Test 3, Day 14) (Guastella *et al*, 2013; MacDonald *et al*, 2011), the reported results are likely to be driven by oxytocin impact at the time of administration (Test 2, ie, during memory encoding). One possible option is that oxytocin ability to motivate affiliative, trusting and explorative behavior (Bartz *et al*, 2011; Feldman, 2012; Kosfeld *et al*, 2005; Meyer-Lindenberg *et al*, 2011) may create a response bias that eliminates the tendency for ratiocination in the face of contradictory social information during Test 2. The immediate result could be quick and heuristic processing, but diminished time and resource allocation necessary for deep encoding and consolidation of long-term memories (Cowan, 1995; Dudai, 2002).

Our results demonstrate that under some circumstances oxytocin administration may induce an immediate behavioral effect that does not involve a change to internal representations. Such an effect will be transient (Asch, 1952; Kelman, 1961; Smith and Mackie, 2007), and most likely not persist in the face of environmental changes typical of the human interpersonal environment (Byrne and Whiten, 1989). This distinction may help explain why early extensive animal and human work found a decrease in memory performance associated with oxytocin (Bruins *et al*, 1992; de Wied *et al*, 1993), while recent studies report mixed results or even an improvement in memory (Meyer-Lindenberg *et al*, 2011; Neumann, 2008; Rimmele *et al*, 2009; Striepens *et al*, 2011).

Future studies, using complementary methodologies, should attempt to further elucidate the multi-level brain underpinnings of the effects described here (Nyberg, 2014). For example, it is possible that oxytocin induces a decrease in amygdala activity (Domes *et al*, 2007; Striepens *et al*, 2011), which has been previously found to subservise encoding of persistent memory errors in this task (Edelson *et al*, 2011). This in turn may inhibit amygdalo-hippocampal and

amygdalo-striatal output pathways related to other neurotransmitters found to affect social learning and trustworthiness such as serotonin and dopamine (Simonsen *et al*, 2014; Skuse and Gallagher, 2011).

Several caveats should be noted. Due to ethical and scientific reasons, our sample population included only male participants (van IJzendoorn and Bakermans-Kranenburg, 2012; MacDonald *et al*, 2011; Meyer-Lindenberg *et al*, 2011) (see Supplementary Information for more details). Thus, conclusions refer to a male population. Secondly, we tested healthy volunteers. Given the different brain morphology and reactivity to pharmacological agents of abnormal populations, our results need to be verified separately in psychopathological populations. In this regard, it is worth noting that processes that may contribute to the observed pattern of results, such as social motivation and exploration, as well as concern for group norms, may be affected in psychopathologies such as autism and schizophrenia (Abu-Akel *et al*, 2014; Churchland and Winkielman, 2012; Meyer-Lindenberg *et al*, 2011; Yafai *et al*, 2013). Thirdly, in accordance with the majority of studies using intranasal oxytocin (Guastella *et al*, 2013; MacDonald *et al*, 2011), we administered a single dose of 24 IU. It is possible that continuous administration over longer time periods or at different dosage would have different effects. Finally, despite the fact that memory modifications constitute a powerful example of internalization of social information, it is not the only measure of persistent social effects. Other measures, such as long-lasting changes in preferences, might additionally be studied.

In conclusion, our results suggest an oxytocin-mediated system that enhances the immediate effects of the social milieu on behavior, but concomitantly reduces the internalization of social influence. Abnormal activity of this system may result in untoward phenomena such as excessive “groupthink” behavior (Janis, 1972) and possibly contribute to psychopathological symptoms. More generally, we believe that the subsumption of long-term testing may lead to the optimization of clinical protocols and to enhanced effectivity in oxytocin-based treatments (Feifel, 2012; Harris and Carter, 2013).

## FUNDING AND DISCLOSURE

All authors (MGE, MS, AW, SY, TS, and YD) declare no biomedical financial interests or potential conflicts of interest. MGE and YD are supported by a Weizmann Institute-UK Making connections Grant. TS is supported by a Wellcome Trust Career Development Fellowship. YD is supported by the Center of Research Excellence in the Cognitive Sciences (I-CORE) and by the Human Brain Project of the EC.

## ACKNOWLEDGEMENTS

We thank R. Feldman for advice and comments on previous drafts, Y. Benjamini for advice on statistics, and A. Ben-Yakov, N. Cohen, J.G. Edelson, A. Mendelsohn, R. Paz, P. Riva, E. Schneidman, and S. Shamay-Tsoory for their comments.

## REFERENCES

- Abu-Akel A, Fischer-Shofty M, Levkovitz Y, Decety J, Shamay-Tsoory S (2014). The role of oxytocin in empathy to the pain of conflictual out-group members among patients with schizophrenia. *Psychol Med* 1–10 (doi:10.1017/S003329171400097X).
- Allen V (1975). Social support for nonconformity. *Adv Exp Soc Psychol* 8: 1–43.
- Asch SE (1952). Group forces in the modification and distortion of judgments. *Soc Psychol (Gott)* 649: 450–501.
- Bales KL, Perkeybile AM, Conley OG, Lee MH, Guoynes CD, Downing GM *et al* (2013). Chronic intranasal oxytocin causes long-term impairments in partner preference formation in male prairie voles. *Biol Psychiatry* 74: 180–188.
- Bartz J, Zaki J, Bolger N, Ochsner K (2011). Social effects of oxytocin in humans: context and person matter. *Trends Cogn Sci* 15: 301–309.
- Baumeister RF (1982). A self-presentational view of social phenomena. *Psychol Bull* 91: 3–26.
- Benjamini Y, Heller R, Yekutieli D (2009). Selective inference in complex research. *Phil Trans R Soc A* 367: 4237–4253.
- Boyd R, Richerson PJ, Richardson PJ (1988). An evolutionary model of social learning: the effects of spatial and temporal variation. *Soc Learn Psychol Biol Perspect* 29–48.
- 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.
- Byrne R, Whiten A (1989). *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans*. Oxford University Press: USA.
- Carter CS (2013). Oxytocin pathways and the evolution of human behavior. *Annu Rev Psychol* 65: 17–39.
- Churchland PS, Winkielman P (2012). Modulating social behavior with oxytocin: how does it work? What does it mean? *Horm Behav* 61: 392–399.
- Cohen J (1988). *Statistical power analysis for the behavioral sciences*. Erlbaum: Hillsdale, New Jersey, USA.
- Cowan N (1995). *Attention and memory: An integrated framework*. Oxford University Press: Oxford, UK.
- Cumming G (2013). The New Statistics: Why and How. *Psychol Sci* 25: 7–29.
- Deutsch M, Gerard HB (1955). A study of normative and informational social influences upon individual judgment. *J Abnorm Psychol* 51: 629–636.
- Domes G, Heinrichs M, Gläscher J, Büchel C, Braus D, Herpertz S (2007). Oxytocin attenuates amygdala responses to emotional faces regardless of valence. *Biol Psychiatry* 62: 1187–1190.
- Dudai Y (2002). *Memory from A to Z: Keywords, concepts, and beyond*. Oxford University Press: Oxford, UK.
- Edelson M, Sharot T, Dolan RJ, Dudai Y (2011). Following the crowd: brain substrates of long-term memory conformity. *Science* 333: 108–111.
- Edelson MG, Dudai Y, Dolan RJ, Sharot T (2014). Brain substrates of recovery from misleading influence. *J Neurosci* 34: 7744–7753.
- Efron B, Tibshirani R (1993). *An introduction to the bootstrap*. Chapman & Hall: Boca Raton, FL, USA.
- Feifel D (2012). Oxytocin as a potential therapeutic target for schizophrenia and other neuropsychiatric conditions. *Neuropsychopharmacology* 37: 304–305.
- Feldman R (2012). Oxytocin and social affiliation in humans. *Horm Behav* 61: 380–391.
- Geenen V, Adam F, Baro V, Mantanus H, Ansseau M, Timsit-Berthier M *et al* (1988). Inhibitory influence of oxytocin infusion on contingent negative variation and some memory tasks in normal men. *Psychoneuroendocrinology* 13: 367–375.
- Guastella AJ, Hickie IB, McGuinness MM, Otis M, Woods EA, Disinger HM *et al* (2013). Recommendations for the standardisation

- of oxytocin nasal administration and guidelines for its reporting in human research. *Psychoneuroendocrinology* **38**: 612–625.
- Harris JC, Carter CS (2013). Therapeutic interventions with oxytocin: current status and concerns. *J Am Acad Child Adolesc Psychiatry* **52**: 998–1000.
- Haun DBM, van Leeuwen EJ, Edelson MG (2013). Majority influence in children and other animals. *Dev Cogn Neurosci* **3**: 61–71.
- Hayes A (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press: New York, USA.
- Heinrichs M, Domes G (2008). Neuropeptides and social behaviour: effects of oxytocin and vasopressin in humans. *Prog Brain Res* **170**: 337–350.
- Hollander E, Bartz J, Chaplin W, Phillips A, Sumner J, Soorya L et al (2007). Oxytocin increases retention of social cognition in autism. *Biol Psychiatry* **61**: 498–503.
- Huang H, Michetti C, Busnelli M, Managò F, Sannino S, Scheggia D et al (2013). Chronic and acute intranasal oxytocin produce divergent social effects in mice. *Neuropsychopharmacology* **39**: 1102–1114.
- Izuma K (2013). The neural basis of social influence and attitude change. *Curr Opin Neurobiol* **23**: 456–462.
- Izuma K, Adolphs R (2013). Social manipulation of preference in the human brain. *Neuron* **78**: 563–573.
- Janis IL (1972). *Victims of groupthink: A psychological study of foreign-policy decisions and fiascoes*. Houghton Mifflin: Oxford, UK, 277.
- Kelman HC (1961). Processes of opinion change. *Public Opin Q* **25**: 57.
- Kosfeld M, Heinrichs M, Zak P, Fischbacher U, Fehr E (2005). Oxytocin increases trust in humans. *Nature* **435**: 673–676.
- Lakens D (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front Psychol* **4**: 863.
- Lewandowsky S, Ecker UKH, Seifert CM, Schwarz N, Cook J (2012). Misinformation and its correction: continued influence and successful debiasing. *Psychol Sci* **13**: 106–131.
- MacDonald E, Dadds M, Brennan J, Williams K, Levy F, Cauchi A (2011). A review of safety, side-effects and subjective reactions to intranasal oxytocin in human research. *Psychoneuroendocrinology* **36**: 1114–1126.
- MacKinnon DP, Fairchild AJ, Fritz MS (2007). Mediation analysis. *Annu Rev Psychol* **58**: 593–614.
- Marsella AJ (1975). Conformity and psychopathology: a comparative study of conformity behaviors in manic-depressive, paranoid schizophrenic and normal populations. *J Clin Psychol* **31**: 402–408.
- Meade ML, Roediger HL (2002). Explorations in the social contagion of memory. *Mem Cogn* **30**: 995–1009.
- Meyer-Lindenberg A, Domes G, Kirsch P, Heinrichs M (2011). Oxytocin and vasopressin in the human brain: social neuropeptides for translational medicine. *Nat rev Neurosci* **12**: 524–538.
- Moreland RL, Levine JM (1989). Newcomers and oldtimers in small groups. *Psychol Gr Infl* **143**–186.
- Neumann ID (2008). Brain oxytocin: a key regulator of emotional and social behaviours in both females and males. *J Neuroendocr* **20**: 858–865.
- Nyberg F (2014). Structural plasticity of the brain to psychostimulant use. *Neuropharmacology* (doi:10.1016/j.neuropharm.2014.07.004; e-pub ahead of print).
- Ortmann A, Hertwig R (2002). The costs of deception: evidence from psychology. *Exp Econ* **5**: 111–131.
- Owen J, Hauth JC (1996). Concentrated oxytocin plus low-dose prostaglandin E2 compared with prostaglandin E2 vaginal suppositories for second-trimester pregnancy termination. *Obs Gynecol* **88**: 110–113.
- Preacher K, Kelley K (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychol Methods* **16**: 93–115.
- Rhode M (2012). Whose memories are they and where do they go? Problems surrounding internalization in children on the autistic spectrum. *Int J Psychoanal* **93**: 355–376.
- Rimmele U, Hediger K, Heinrichs M, Klaver P (2009). Oxytocin makes a face in memory familiar. *J Neurosci* **1**: 38–42.
- Rothbaum F, Weisz JR, Snyder SS (1982). Changing the world and changing the self: A two-process model of perceived control. *J Pers Soc Psychol* **42**: 5–37.
- Shalvi S, De Dreu CKW (2014). Oxytocin promotes group-serving dishonesty. *Proc Natl Acad Sci USA* **111**: 5503–5507.
- Simonsen A, Scheel-Krüger J, Jensen M, Roepstorff A, Møller A, Frith CD et al (2014). Serotonergic effects on judgments and social learning of trustworthiness. *Psychopharmacology (Berl)* **231**: 2759–2769.
- Skuse DH, Gallagher L (2011). Genetic influences on social cognition. *Pediatr Res* **69**: 85R–91R.
- Smith E, Mackie D (2007). *Social psychology*. Psychology Press: London, UK.
- Stallen M, De Dreu C, Shalvi S, Smidts A, Sanfey A (2012). The herding hormone oxytocin stimulates in-group conformity. *Psychol Sci* **23**: 1288–1292.
- Stang DDJ (1976). Ineffective deception in conformity research: Some causes and consequences. *Eur J Soc Psychol* **6**: 353–367.
- Striepens N, Kendrick KM, Maier W, Hurlmann R (2011). Prosocial effects of oxytocin and clinical evidence for its therapeutic potential. *Front Neuroendocr* **32**: 426–450.
- van IJzendoorn MH, Bakermans-Kranenburg M (2012). A sniff of trust: meta-analysis of the effects of intranasal oxytocin administration on face recognition, trust to in-group, and trust to out-group. *Psychoneuroendocrinology* **37**: 438–443.
- Wied D de, Diamant M, Fodor M (1993). Central nervous system effects of the neurohypophyseal hormones and related peptides. *Front Neuroendocr* **14**: 251–302.
- Yafai A-F, Verrier D, Reidy L (2013). Social conformity and autism spectrum disorder: A child-friendly take on a classic study. *Autism* **18**: 1007–1013.
- Young LJ, Wang Z (2004). The neurobiology of pair bonding. *Nat Neurosci* **7**: 1048–1054.

Supplementary Information accompanies the paper on the Neuropsychopharmacology website (<http://www.nature.com/npp>)