PROCEEDINGS B

rspb.royalsocietypublishing.org

Research

Cite this article: McClung JS, Triki Z, Clément F, Bangerter A, Bshary R. 2018 Endogenous oxytocin predicts helping and conversation as a function of group membership. *Proc. R. Soc. B* **285**: 20180939.
http://dx.doi.org/10.1098/rspb.2018.0939

Received: 26 April 2018
Accepted: 8 June 2018

Subject Category:
Behaviour

Subject Areas:
behaviour, physiology

Keywords:
oxytocin, cooperation, group membership, helping behaviour

Author for correspondence:
Jennifer Susan McClung
e-mail: jennifer.mcclung@unine.ch

Endogenous oxytocin predicts helping and conversation as a function of group membership

Jennifer Susan McClung[†], Zegni Triki[†], Fabrice Clément, Adrian Bangerter and Redouan Bshary

Centre for Cognitive Science, Institute of Biology, University of Neuchâtel, Emile-Argand 11, 2000 Neuchâtel, Switzerland

JSM, 0000-0002-3502-4120; RB, 0000-0001-7198-8472

Humans cooperate with unrelated individuals to an extent that far outstrips any other species. We also display extreme variation in decisions about whether to cooperate or not, and the mechanisms driving this variation remain an open question across the behavioural sciences. One candidate mechanism underlying this variation in cooperation is the evolutionary ancient neurohormone oxytocin (OT). As current research focuses on artificial administration of OT in asocial tasks, little is known about how the hormone in its naturally occurring state actually impacts behaviour in social interactions. Using a new optimal foraging paradigm, the 'egg hunt', we assessed the association of endogenous OT with helping behaviour and conversation. We manipulated players' group membership relative to each other prior to an egg hunt, during which they had repeated opportunities to spontaneously help each other. Results show that endogenous baseline OT predicted helping and conversation type, but crucially as a function of group membership. Higher baseline OT predicted increased helping but only between in-group players, as well as decreased discussion about individuals' goals between in-group players but conversely more of such discussion between out-group players. Subsequently, behaviour but not conversation during the hunt predicted change in OT, in that out-group members who did not help showed a decrease in OT from baseline levels. In sum, endogenous OT predicts helping behaviour and conversation, importantly as a function of group membership, and this effect occurs in parallel to uniquely human cognitive processes.

1. Introduction

One major point of discussion between researchers of human behaviour has been the ultimate and proximate mechanisms that drive our unmatched cooperation between unrelated individuals. On an ultimate level, cultural group selection [1,2], cooperative breeding [3], reputation [4] and high levels of partner choice [5] may all interact to promote human cooperation. However, it is increasingly clear that to understand the vast variation humans display in their cooperative behaviour [6–8], we need to understand both the psychological and physiological mechanisms underlying decision-making in a cooperative context.

Regarding psychological mechanisms, group membership has emerged as an influential social factor in producing variation in cooperation. Decades of psychological and economic studies, both experimental and in the field, have shown that we have a bias to cooperate more with members of our own group, even at personal cost, compared to members of other groups (see [9] for a meta-analysis). Driving this bias, shared intentionality, or the ability to mentally represent another's goals and then to assume them as one's own [10], is a psychological mechanism that increases cooperation between in-group members [11]. Physiologically, the evolutionary ancient neuropeptide oxytocin (OT), which is produced in the hypothalamus in response to social

---

[†] These authors contributed equally to this work.

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.4135313.

THE ROYAL SOCIETY
PUBLISHING

bonding and affiliation, also seems to support an in-group bias in cooperation [12]. OT has also been co-opted to regulate group living, specifically by facilitating compliance with group norms and increasing trust, empathy and cooperation between members of the same group (for a review, see [13]).

However, the conclusions about the role OT plays in human social life largely stem from research that uses administration of the hormone via a nasal spray. The effects of exogenous OT are not consistent as individual and social factors moderate its pro-social effects. For example, individuals with borderline personality disorder playing an assurance game make less trusting decisions when given exogenous OT [14]. Regarding social factors, group membership interacts with exogenous OT to impact cooperation [15]. Using an inter-group Prisoner's Dilemma-maximizing-differences game in which players could choose between donating money to benefit in-group members only or to benefit in-group members while harming the out-group (by lowering the out-group's rewards), De Dreu *et al*. [16] showed that exogenous OT increased donations beneficial to the in-group but did not affect donations that would also cause out-group harm. In a follow-up using a similarly structured game, males given OT cooperated less with out-group members, but specifically when they themselves or their own in-group was more vulnerable. As such, exogenous OT seems to increase bias in favour of the in-group while having less of an impact on bias against the out-group.

Our understanding of how OT actually functions in social interactions is currently limited by the fact that exogenous application is artificial and as such can mask or interact with naturally occurring levels of the hormone (see [17] for a review). That is, while exogenous application can inform us as to the explicitly causal role of OT in behaviour, this procedure can offer no explanation about the actual role the hormone plays in its natural state as application necessarily alters baseline levels. Of the few studies that have looked at endogenous OT in adult interactions, the norm has been to assess the impact of pro-social behaviour on subsequent OT levels. For instance, trust games affect endogenous OT levels: sharing and receiving secrets increases OT levels from baseline [18], as does being the recipient of an intentional compared to unintentional monetary donation [19]. One study by Zhong *et al*. [20] looked at the effect of baseline OT on subsequent behaviour and reported that both very high and very low levels of endogenous OT correlated with increased monetary donations in a standard trust game, but only in males. Another study by Levy *et al*. [21] examined baseline OT in relation to in-group bias and found that OT correlated with increased in-group bias, but only in participants that represented a majority group (Jewish–Israeli participants) and not in those representing a minority group (Arab–Palestinian participants).

While the emerging picture from research on both exogenous and endogenous OT suggests that OT is indeed involved in modulating cooperation, all studies to date suffer from a peculiar caveat: subjects never experienced proper face-to-face social interactions but interacted via written messages, a computer interface, or not at all. As such, we can only guess how naturally occurring OT would impact behaviour and conversation in the social situations that it has supposedly evolved to facilitate [12]. Given the aforementioned results, we hypothesize that endogenous OT does impact behaviour in actual social interactions, specifically cooperative behaviour, including the conversation that would facilitate it, but that crucially this effect is likely to be a function of group membership. It is likely that increased baseline levels of endogenous OT will be associated with more cooperation between in-group members than out-group members. An open question is the impact that increased cooperation may have on subsequent levels of OT, regardless of group membership. Given the role that communication via language can play in increasing in-group cooperation, we would also expect higher baseline OT to be associated with increased cooperative conversation, specifically conversation relating to the psychological mechanism driving variation in cooperation, i.e. shared intentionality [11]. As mentioned above, shared intentionality is the ability to represent and assume another's goals, and may be manifest in linguistic exchange between potential cooperators. As this is the psychological mechanism via which in-group members increase their cooperation with each other relative to out-group members, we also expected that OT would predict shared intentionality in conversation.

To examine these hypotheses, we designed a task that is loosely modelled on the Prisoner's Dilemma. In our task, the 'egg hunt', pairs of players engaged in a 5 min hunt for screws of three different colours (only two of which were rewarded) that were hidden in small, plastic eggs throughout a room. Each player was rewarded for the total collection of screws of one assigned colour (either red or blue, with green screws being irrelevant to the total reward). Therefore, players could choose to help the other during the hunt each time they came across a screw of the other's colour. Crucially, this option was not mentioned in the instructions and players therefore had to discover spontaneously that the possibility of mutual helping existed. As such, helping was costly (in terms of time taken away from the collection of one's own screws), which meant that it only paid to help on the condition that the other player helped conditionally as well, whereas it did not pay to help if the partner followed an unconditional strategy (e.g. always helping or always neglecting to help). The experiment aimed to isolate both cognitive and physiological mechanisms that may produce variation in levels of helping. We focused on the impact of two social factors, group membership and communication. This was because, as discussed above, shared group membership is known to increase cooperation, and because one of our aims was to examine cooperation in a natural social setting, hence our inclusion of normal linguistic exchange. Using a 2 × 2 design, we either allowed participants to talk or prohibited talking, and we formed in-group or out-group pairs based on trivial criterion, similarly to the groups used by De Dreu *et al*. [15]. Already published data from the egg hunt experiment showed that, overall, talking increased cooperation, and in terms of psychological mechanisms, variation in cooperation was driven by an increase in shared intentionality talk between in-group members over out-group members [11]. The aim of the current study was to examine the degree to which endogenous OT levels may also explain variation in cooperation that stems from group membership, as well as how variation in helping may affect OT levels.

To achieve these goals, we assessed whether endogenous baseline levels of OT may predict variation in cooperation and the types of conversation that drive it, and whether cooperation or types of conversation can in turn affect endogenous levels of OT. As OT in saliva correlates positively with circulating OT [22], we measured OT from saliva samples taken before and after the egg hunt. Given the aforementioned effect of exogenous OT application on cooperation between in-group members [15], we predicted that higher baseline OT would be associated with increased cooperation and shared intentionality (displayed in conversations) between in-group members. We also expected that increased cooperation and shared intentionality during the hunt would lead to an increase in endogenous OT from baseline levels.

2. Material and methods

(a) Participants

The University of Neuchâtel's ethics committee approved the current study. In total, 138 healthy participants took part in the present study: 36 males and 102 females. This formed a total of 69 pairs of players who completed the experiment in tandem. Players were assigned to pairs of the same gender and it was confirmed that pairs did not know each other prior to the experiment. All pairs were tested in a single session lasting approximately 40 min. Experiments were conducted in French. All players were naive to the experimental hypothesis, told that their participation was entirely voluntary and that they could leave at any time. They were also informed that their data would be treated confidentially and used anonymously in publication. All players gave informed consent, and were fully debriefed at the end of each experiment. All players earned CHF 10 (Swiss Francs) as a show-up fee and up to CHF 20 as a top-up payment as a function of their performance in the experiment.

(b) Procedure

The experiment was conducted in the Biology Department at the University of Neuchâtel. For our 2 × 2 design, group membership × talking condition, we operationalized group membership by categorizing players as either in-group or out-group relative to each other after they filled in a 10-item questionnaire about their food preferences (see electronic supplementary material for complete Food Preferences Questionnaire). After completion of the questionnaire but before categorization, saliva samples were collected to assess baseline OT levels (see electronic supplementary material for Experiment timeline, figure S1). Saliva samples were taken using cotton swabs (Salivettes®, Sarstedt, Germany) that participants were asked to chew for 2 min. Players were then arbitrarily designated as either 'apple' or 'orange', without any claim of a causal relationship between their questionnaire responses and this categorization. This was done to conform to standards in economics research that prohibit explicit deception. Players were then asked to wear laboratory coats with the corresponding colour and logo (green for 'apples' and orange for 'oranges'). This was to make categorization salient to players and form the 'In-group' condition (two 'apples' or two 'oranges') and the 'Out-group' condition (one 'apple' and one 'orange'). After categorization, we gave participants instructions for the egg hunt, during which we operationalized talking by simply instructing players that they were either allowed to talk freely ('Talking') or prohibited from talking during the hunt ('Not talking'; see electronic supplementary material for Instructions and experimental set-up, figure S2).

The instructions informed players that they would take part in an egg hunt, and that each egg had either a red, blue or green screw within. Players were given a board into which they could screw the screws they collected and they were told that one of them would be rewarded for all the red screws collected (at CHF 1 each), one would be rewarded for all the blue screws collected (also at CHF 1 each), and that the green screws were not rewarded. Players were told that they could carry out the hunt in any way that allowed them to maximize their monetary reward. Importantly, they were never explicitly told they could collect the other's screws or cooperate in any way but were instead told that their reward would be based on the number of screws 'screwed into a board' at the end of the hunt without any reference to which board the screws were required to be screwed into. Players were then told whether they would be allowed to talk or not and they were given a 1 min 'countdown' before the hunt started. Players were given 5 min for the egg hunt. Behavioural data were recorded using eye-tracking glasses (ETG 2.1 models provided by SensoMotoric Instruments GmbH, Germany). Immediately at the end of the hunt, a second saliva sample was collected to assess the final OT levels using the same method described above. The time lapse between the first and second saliva sample was approximately 11 min. All saliva samples were frozen at −20°C for further hormonal analysis.

(c) Behaviour analysis

In order to assess the interaction between OT and cooperative behaviours, we used previously published data on the cooperative behaviour of players during the egg hunt [11]. In that study, individuals were assessed for levels of cooperation during the hunt by categorizing each decision made about a partner's screw as either costly helping, no-cost helping, or neglect (see electronic supplementary material for complete Behaviour Coding Scheme). Costly helping cost the player in terms of time on their own hunt (i.e. carrying or screwing the other's screw into one's own board), whereas no-cost helping cost nothing in terms of time (i.e. leaving the other's screw outside the egg). Neglect constituted simply leaving the other's screw in the egg as they had found it. For the purposes of this study, we focused on costly helping because it is specifically helping at a cost to oneself that shared group membership increases [11], an effect also seen under artificial external application of OT [23]. For our statistical analyses, a binary response variable was generated from the proportion of helping from the total screws of the other's colour found, where the proportions greater than 0.5 were scored 'yes' and proportions less than 0.5 were scored 'no'.

(d) Conversation analysis

In order to assess the interaction of OT and cognitive processes as evidenced via conversation, we used data from the transcribed conversations of players in the talking condition in the egg hunt talk [11]. For that study, conversations were segmented into utterances and each was categorized as one of the following: (i) shared intentionality talk, (ii) individual goal talk, (iii) task talk, or (iv) other (see electronic supplementary material for complete Conversation Coding Scheme). (i) Shared intentionality talk consisted of any reference to the hunt in terms of a shared or common goal, which included both planning the shared goal before the hunt (e.g. 'we should collect both colours and trade at the end') and referencing the shared goals during the hunt (e.g. 'I've got one of yours'). (ii) Individual goal talk consisted of any reference to the task in terms of distinctly separate goals that were never shared (e.g. 'you collect your reds and I'll get the blues'). (iii) Task talk consisted of any reference to the practical aspects of the task that involved neither shared nor individual goals (e.g. 'how much time do we have left?'). (iv) Other talk consisted of all utterances that did not fall into the above categories (e.g. 'I am from Neuchâtel').

(e) Oxytocin analysis

The Salivettes® were stored immediately on ice until the end of each trial. Between 0.4 and 1 ml of saliva collected via centrifugation was necessary for analysis to be performed. Salivary OT was first extracted and then analysed in the laboratory at the University of Neuchâtel using the enzyme immunoassay technique ELISA (see electronic supplementary material for full details of extraction and ELISA procedures). Owing to missing samples and readings that did not fulfil quality requirements in the ELISA process, we obtained matched samples (both baseline OT and final OT) for 47 participants. (Analyses on both datasets—the dataset with only matched samples and the dataset with some unmatched samples—are presented in the electronic supplementary material for the sake of comparison, and we show similar results for both sets.) All analyses here focus on the 47 participants for whom we had matched samples so as to avoid possible bias from participants from whom we only had one sample. The results of all models used to assess first the relation of baseline OT with helping behaviour and conversation, and second, helping behaviour and conversation on change in OT, are summarized in table 1.

(f) Data analyses

We used R v. 3.5.0 (2018-04-23) for all statistical analysis (see electronic supplementary material for all analysis details).

3. Results

(a) Oxytocin and behaviour

First, we assessed the relationship between baseline OT levels and cooperative behaviour. We found an interaction between group membership and baseline OT levels (generalized linear model, GLM: $X^2 = 7.190$, $p = 0.007$, figure 1*a*); that is, increased baseline OT predicted costly helping but only in the in-group: out-group members' costly helping was unaffected by baseline OT. We found no interaction between baseline OT and gender on costly helping.

Overall change in OT throughout the experiment was predicted by an interaction of costly helping with group membership (generalized least squares GLS: $X^2 = 4.833$, $p = 0.027$, figure 1*b*). Post hoc analyses showed that of all participants who did not help, out-group members showed a significantly greater decrease from baseline OT compared to in-group members ($p = 0.006$). Of all participants who helped, there was no difference between in-group and out-group members ($p = 0.414$). There were also no differences within the out-group between helpers and non-helpers ($p = 0.374$) or within the in-group between helpers and non-helpers ($p = 0.387$). Further post hoc analyses followed the model, where we compared each interaction level to the null hypothesis, or no change in the OT levels. We found that out-group members that did not help had a significant decrease in the OT levels from baseline ($p = 0.020$). However, no significant changes from baseline OT were found in any of the other conditions (out-group members that helped: $p = 0.805$; in-group members that helped: $p = 0.805$, in-group members that did not help: $p = 0.253$).

(b) Oxytocin and conversation

When analysing the content of participants' conversations in relation to OT, we did not find an effect of initial OT on shared intentionality talk as hypothesized. That is, there was no main effect of baseline OT in predicting the amount of shared intentionality talk players engaged in (linear model, LM: $F\ (1, 19) = 1.560$, $p = 0.227$), nor did we see an interaction of baseline OT with group membership on shared intentionality talk (LM: $F\ (1, 19) = 0.119$, $p = 0.734$; figure 2*a*). We did, however, find a significant interaction between group membership and baseline OT on individual goal talk (LM: $F\ (1, 20) = 8.920$, $p = 0.0073$; figure 2*b*). That is, higher baseline OT levels predicted more individual goal talk between out-group members but less between in-group members. We found no interaction of initial OT with group membership on the other types of talk (task talk: LM: $F\ (1, 20) = 0.0347$, $p = 0.854$, figure 2*c*; other talk: LM: $F\ (1, 20) = 2.693$, $p = 0.116$, figure 2*d*), although we did find an interaction of task talk and gender (LM: $F\ (1, 20) = 27.832$, $p < 0.001$). We found no main effects of initial OT on other types of talk (task talk: LM: $F\ (1, 20) = 1.070$, $p = 0.313$; other talk: LM: $F\ (1, 20) = 0.154$, $p = 0.699$).

However, no type of conversation predicted overall change in OT throughout the experiment as a function of group membership (shared intentionality talk: LM: $F\ (1, 19) = 0.562$, $p = 0.463$, figure 3*a*; individual goal talk: LM: $F\ (1, 20) = 1.758$, $p = 0.199$; figure 3*b*; task talk: $F\ (1, 19) = 1.038$, $p = 0.321$, figure 3*c*; other talk: $F\ (1, 19) = 0.101$, $p = 0.753$, figure 3*d*).

4. Discussion

The egg hunt paradigm is a particular form of an iterated Prisoner's Dilemma game as players may, in essence, choose to view their goals as either individual or shared, with the latter potentially leading to helping at a small cost. This means that failure to help is not perceived as defection as in the classic Prisoner's Dilemma. Using this paradigm, we were first able to gauge the relationship of baseline levels of endogenous OT on spontaneous helping and conversation during the experiment, and then the association between spontaneous helping and conversation with change in OT from baseline to final samples.

(a) Endogenous oxytocin and cooperation

First, we found that baseline OT levels predicted cooperative behaviour during the hunt, but importantly this effect was a function of group membership. Baseline OT predicted only in-group members' helping, with higher baseline OT predicting more costly helping, whereas out-group members' costly helping was not affected by baseline OT. These results confirm that which results from studies using exogenous OT can only suggest: that spontaneous helping is promoted by naturally occurring levels of OT, specifically helping that costs the individual but benefits the in-group [23]. For instance, De Dreu *et al.* [15] showed that artificially administered OT can induce higher monetary donations to in-group members compared to out-group members at a cost to the individual donator. Our results parallel the De Dreu study and suggest that baseline OT is a mechanism that mediates spontaneous helping between in-group members. This mechanism may have older evolutionary roots as similar results have recently been found in chimpanzees: animals preparing for inter-group conflict, and as such requiring coordination and cohesion within their own in-groups, have elevated endogenous OT compared to individuals engaged in other in-group interactions that do not involve conflict with an out-group [24]. As mentioned above, few studies have looked at endogenous OT (e.g. [18,19]). To our knowledge, only one study has examined the effects of naturally occurring baseline OT on subsequent behaviour,

Table 1. Data analysis outcomes for all models. Predictors of interest are reported in the table. Significant effects (threshold set at $\alpha = 0.05$, $p \leq 0.05$) are indicated in bold.

	model class	response variable	explanatory variable	χ^2/*F*-value	d.f.	*p*-value	sample size *n*	explained deviance	explained variance from adjusted R^2	figure
baseline OT levels	GLM	cooperation as costly helping	baseline OT	1.766	1	0.183	47	**29.10%**		—
			baseline OT × group membership	**7.190**	**1**	**0.007**	47			figure 1*a*
			baseline OT × gender	2.311	1	0.128	47			—
	LM	shared intentionality talk	baseline OT	1.560	1	0.227	26	—	28.56%	—
			baseline OT × group membership	0.119	1	0.734	26			figure 2*a*
			baseline OT × gender	3.213	1	0.088	26			—
	LM	individual goal talk	baseline OT	1.178	1	0.290	26	—	**39.46%**	—
			baseline OT × group membership	**8.920**	**1**	**0.007**	26			figure 2*b*
			baseline OT × gender	2.712	1	0.115	26			—
	LM	task talk	baseline OT	1.070	1	0.313	26	—	**52.34%**	—
			baseline OT × group membership	0.0347		0.854				figure 2*c*
			baseline OT × gender	**27.832**		**<0.001**				—
	LM	other talk	baseline OT	0.154	1	0.699	26	—	0.00%	—
			baseline OT × group membership	2.693		0.116				figure 2*d*
			baseline OT × gender	0.890		0.357				—
change in OT levels	GLS	changes in OT levels	**costly helping × group membership**	**4.833**	**1**	**0.027**	47	—	**17.52%**	figure 1*b*
			costly helping × gender	2.629	1	0.104	47			—
	LM	changes in OT levels	shared intentionality × group membership	0.662	1	0.425	26	—	10.01%	figure 3*a*
			shared intentionality × gender	1.698	1	0.208	26			—
	LM	changes in OT levels	individual goal talk × group membership	1.836	1	0.190	26	—	14.31%	figure 3*b*
			individual goal talk × gender	1.764	1	0.199	26			—

(*Continued*.)

Table 1. (*Continued*.)

model class	response variable	explanatory variable	χ^2/*F*-value	d.f.	*p*-value	sample size *n*	explained deviance	explained variance from adjusted R^2	figure
LM	changes in OT levels	task talk × group membership	1.038	1	0.321	26	—	−0.05%	figure 3*c*
		task talk × gender	0.432		0.519	26			—
LM	changes in OT levels	other talk × group membership	0.101	1	0.753	26	—	0.00%	figure 3*d*
		other talk × gender	0.063		0.804	26			—

showing that both very low and very high levels of baseline OT lead to higher monetary donations in a typical trust game [20]. None of these studies looked at OT in relation to behaviour during an actual social interaction, and potentially 'trusting' behaviours were operationalized using the written transmission of secrets [18] or trust games via computer interfaces [19,20]. In contrast with these existing studies, our results using the egg hunt paradigm show that endogenous OT predicts actual social behaviour, specifically spontaneous cooperative decisions between in-group members.

We also expected that cooperative behaviour would be associated with a subsequent increase in endogenous OT, which would possibly be greater in the in-group condition. Contrary to this prediction, few players showed an increase from baseline OT levels, whereas most showed either no change or a decrease. Interestingly, this too was moderated by group membership, in that OT levels decreased significantly from baseline only in out-group members who did not help their partners. We would speculate that, if OT functions to increase the saliency and perception of social cues as well as to sustain affiliative motivation [17], the out-group members who avoided helping each other were already responding appropriately to the social cues of group membership. That is, not helping an out-group member would maintain the social distance instantiated by different social identities, and as such, these players may not have required a further physiological vector to stimulate socially appropriate behaviour. This may have led to the decline in OT. However, such reasoning would also lead us to expect a decrease in OT in in-group members who, conversely, did help each other, and our results do not show such a decrease. It is also possible that our window of time between initial and final samples (11 min) was too small, and that a longer interval would have produced a different pattern of change than what we found. Further work will be necessary to explain this decrease in OT to identify the direction of causality between helping (or its absence) and group membership.

(b) Endogenous oxytocin and conversation

When looking at conversation, we predicted that endogenous OT would predict shared intentionality talk. This is because shared intentionality is the type of conversation that promotes higher levels of cooperation between in-group members [11]. However, we found no association between endogenous OT and shared intentionality, but instead an association between individual goal talk and OT as a function of group membership. This type of talk involved discussing the task in terms of goals that each individual developed and worked towards separately, for example, 'I'll get my blues and you get your reds', and as such required little understanding of or reference to the other's mental states. Interestingly, our results show that higher baseline OT predicted more of this type of talk in the out-group condition but less in the in-group condition. We also predicted that shared intentionality in conversation during the hunt would affect change in baseline levels of endogenous OT. Again, we saw no association of shared intentionality or individual goal talk with change in OT.

These results are somewhat surprising as we would have expected OT to predict the type of talk that promotes helping, shared intentionality, in order to increase the in-group's success. This is especially so as shared intentionality is most closely related to capacities that exogenous OT has been

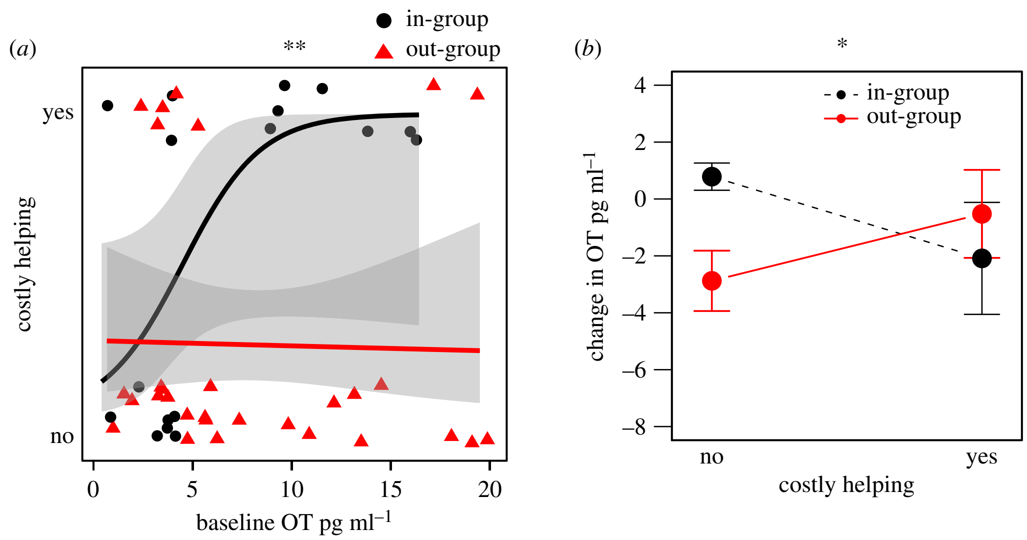

Figure 1. (*a*) Baseline OT levels predicting costly helping as a function of group membership with 95% confidence intervals. Logistic regression lines shown for each condition. (*b*) Costly helping predicting changes in OT levels (from baseline to final sample) as a function of group membership. Means and standard errors of changes in OT levels are shown. $^*p \leq 0.05$, $^{**}p \leq 0.01$.

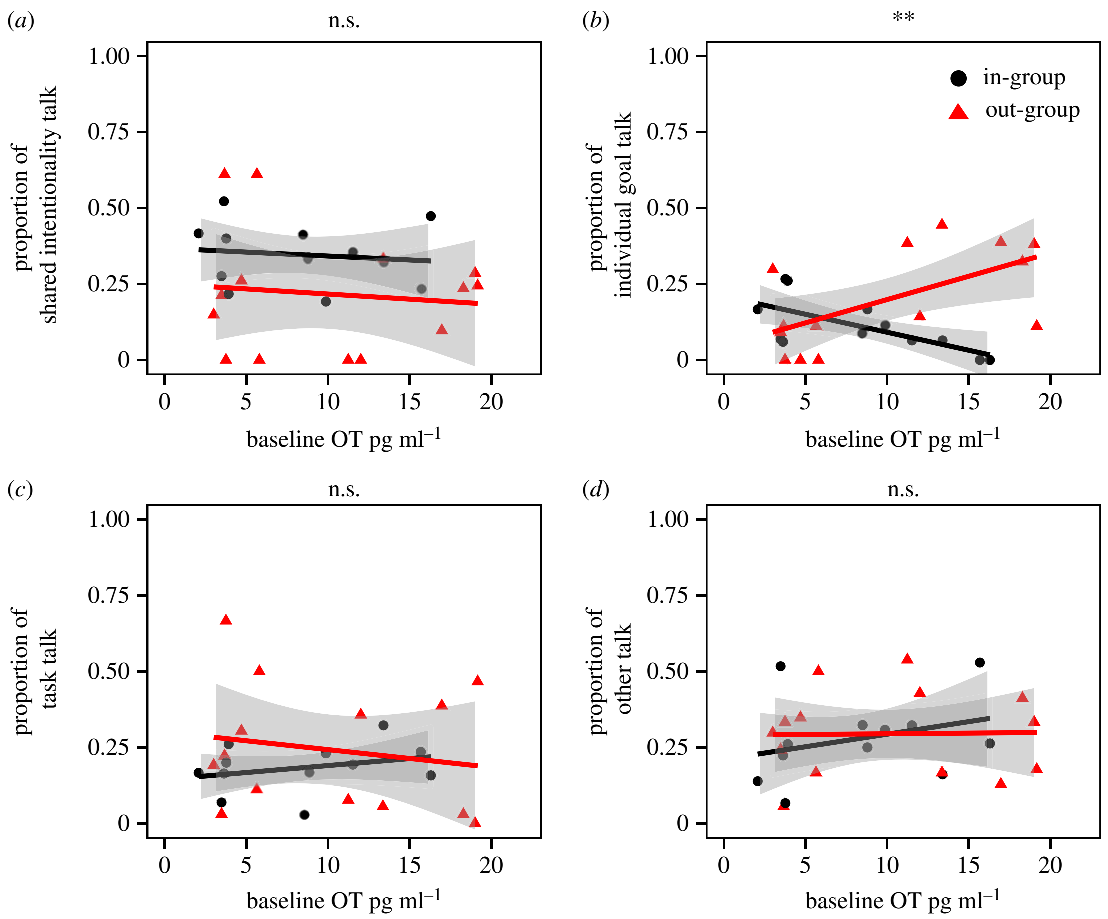

Figure 2. Baseline OT levels predicting different types of talk as a function of group membership with 95% confidence intervals. (*a*) Shared intentionality talk, (*b*) individual goal talk, (*c*) task talk, and (*d*) other talk. Logistic regression lines shown for each condition. $^{**}p \leq 0.01$; n.s., non-significant differences.

shown to impact such as the inference of mental states from pictures of the eyes [25] and discussion of information related to a group's success [26]. So, it could seem odd that in naturalistic social interactions between in-group members like we use here, endogenous OT seems to have no effect on the type of talk that involves understanding and sharing in others' mental states to facilitate cooperation. Instead, the reduction in discussion of individually separate goals predicted by baseline OT between in-group members may already promote helping, as suggested by a preliminary analysis that yielded a negative correlation between individual goal talk and helping (see electronic supplementary material, figure S6). Consequently, we would speculate that these physiological and psychological mechanisms—OT and

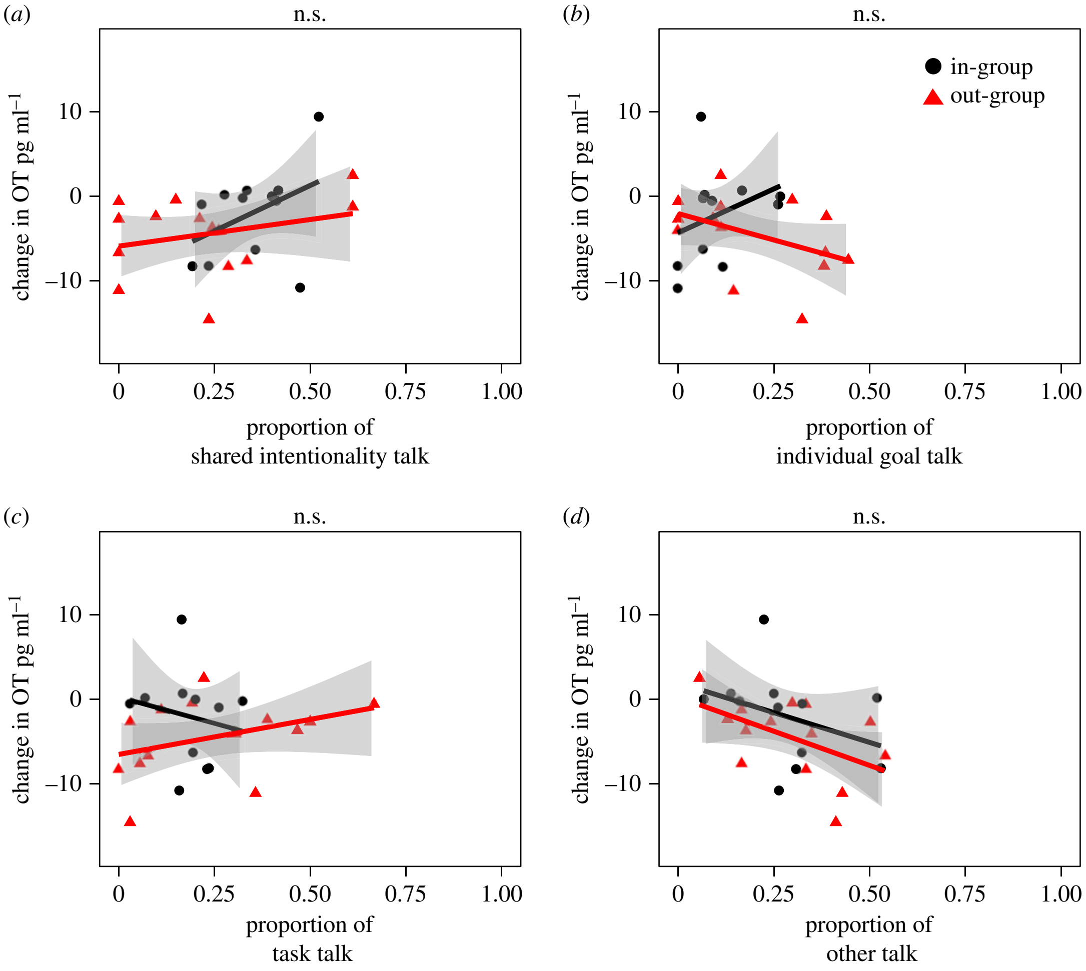

Figure 3. Four different types of talk predicting change in OT (from baseline to final sample) as a function of group membership with 95% confidence intervals. (*a*) Shared intentionality talk, (*b*) individual goal talk, (*c*) task talk, and (*d*) other talk. Linear regression lines shown for each condition. n.s., non-significant differences.

shared intentionality, respectively—are acting in parallel to produce socially appropriate helping behaviour per situation. That is, if baseline OT levels interact with group membership to produce costly helping, it is possibly in these situations that additional shared intentionality talk is not needed to achieve socially appropriate levels of cooperation, and vice versa. In line with this parallel effect hypothesis, it has been proposed that OT produces socially appropriate behaviour via its influence on very simple cognitive processes, such as the motivation to seek social interaction [27], whereas shared intentionality requires much more complex cognition such as representing and interacting with various mental states [28]. If this is the case, then it may follow that OT produces socially appropriate levels of helping (i.e. more in the in-group, less in the out-group) via its impact on the simpler form of cognition. That is, representing another individual's goal will clearly require fewer and simpler cognitive steps than representing another's goal and assuming it as one's own to then negotiate its achievement together as required in shared intentionality. In sum, it seems OT may impact simpler social behaviours, often termed 'naive sociology', while having no impact on social behaviours requiring more complex psychological reasoning [29].

(c) Overall discussion

Taken together, these results shine light on the physiological mechanisms that produce both variation in human cooperation and more broadly allow the maintenance of social biases. First, a huge body of work in social psychology is dedicated to examining the human propensity to form and maintain groups. Favouritism of fellow in-group members over out-group members is well documented (e.g. [30–32]) even in randomly formed groups [33,34]. While the existing literature has shown that exogenous application of OT can produce more helping behaviour with one's own group, our results extend this literature by providing data on how the hormone in its naturally occurring concentrations may impact decisions and behaviour in a natural social interaction.

Our results also add to the literature describing OT as an important neuromodulator in the maintenance of social bonds, even if those bonds result from the arbitrary formation of groups. We provide support for the idea that OT's main role is to improve how people make use of social cues with the aim of producing the most socially appropriate, and therefore useful, behaviour [17]. This is why we see a differential association of baseline OT with behaviour and conversation as a function of group membership, as well as a differential association of behaviour with change in OT as a function of group membership. The variation in players' cooperation as a function of both OT and group membership suggests that OT is a proximate mechanism driving the vast variation we see in human cooperation [1,6], possibly working alongside other psychological mechanisms such as shared intentionality to produce socially appropriate behaviour [11]. Future work will be needed to examine how OT and potentially other physiological mechanisms interact with cognition, specifically whether OT has a differential impact on simple or complex cognitive processes that may be more or less social, to produce variation in cooperative behaviour across social contexts.

Ethics. The study was approved by the ethics committee at the University of Neuchâtel prior to data collection.

Data accessibility. The data are available in the Figshare repository (doi:10.6084/m9.figshare.5620429).

Authors' contributions. J.S.M., R.B., A.B. and F.C. designed the study. J.S.M. collected behavioural data. Z.T. ran the hormonal analyses. J.S.M. and Z.T. analysed the data. J.S.M. and Z.T. wrote the manuscript with input from R.B., F.C. and A.B.

Competing interests. We declare we have no competing interests.

Funding. Funding was provided by the Swiss National Science Foundation (grant no. 31003A_153067/1 to R.B.).

Acknowledgements. We wish to thank Radu Slobodeanu for his assistance with data analysis.

References

1. Fehr E, Gachter S. 2002 Altruistic punishment in humans. *Nature* **415**, 137–140. (doi:10.1038/415137a)
2. Kümmerli R, Burton-Chellew MN, Ross-Gillespie A, West SA. 2010 Resistance to extreme strategies, rather than prosocial preferences, can explain human cooperation in public goods games. *Proc. Natl Acad. Sci. USA* **107**, 10 125–10 130. (doi:10.1073/pnas.1000829107)
3. Burkart JM *et al.* 2014 The evolutionary origin of human hyper-cooperation. *Nat. Commun.* **5**, 4747. (doi:10.1038/ncomms5747)
4. Milinski M, Semmann D, Krambeck H-J. 2002 Reputation helps solve the 'tragedy of the commons'. *Nature* **415**, 424–426. (doi:10.1038/415424a)
5. Raihani NJ, Bshary R. 2015 The reputation of punishers. *Trends Ecol. Evol.* **30**, 98–103. (doi:10.1016/j.tree.2014.12.003)
6. Herrmann B, Thöni C, Gächter S. 2008 Antisocial punishment across societies. *Science* **319**, 1362–1367. (doi:10.1126/science.1153808)
7. Burton-Chellew MN, El Mouden C, West SA. 2016 Conditional cooperation and confusion in public-goods experiments. *Proc. Natl Acad. Sci. USA* **113**, 1291–1296. (doi:10.1073/pnas.1509740113)
8. Baumard N, André J-B, Sperber D. 2013 A mutualistic approach to morality: the evolution of fairness by partner choice. *Behav. Brain Sci.* **36**, 59–78. (doi:10.1017/S0140525X11002202)
9. Balliet D, Wu J, De Dreu CKW. 2014 Ingroup favoritism in cooperation: a meta-analysis. *Psychol. Bull.* **140**, 1556. (doi:10.1037/a0037737)
10. Tomasello M, Carpenter M, Call J, Behne T, Moll H. 2005 Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* **28**, 675–691. (doi:10.1017/S0140525X05000129)
11. McClung JS, Placì S, Bangerter A, Clément F, Bshary R. 2017 The language of cooperation: shared intentionality drives variation in helping as a function of group membership. *Proc. R. Soc. B* **284**, 20171682. (doi:10.1098/rspb.2017.1682)
12. Carter CS. 2014 Oxytocin pathways and the evolution of human behavior. *Annu. Rev. Psychol.* **65**, 17–39. (doi:10.1146/annurev-psych-010213-115110)
13. De Dreu CKW, Kret ME. 2016 Oxytocin conditions intergroup relations through upregulated in-group empathy, cooperation, conformity, and defense. *Biol. Psychiatry* **79**, 165–173. (doi:10.1016/j.biopsych.2015.03.020)
14. Bartz JA, Simeon D, Hamilton H, Kim S, Crystal S, Braun A, Vicens V, Hollander E. 2010 Oxytocin can hinder trust and cooperation in borderline personality disorder. *Soc. Cogn. Affect. Neurosci.* **6**, 556–563. (doi:10.1093/scan/nsq085)
15. De Dreu CKW, Greer LL, Handgraaf MJJ, Shalvi S, Van Kleef GA, Baas M, Ten Velden FS, Van Dijk E, Fieth SWW. 2010 The neuropeptide oxytocin regulates parochial altruism in intergroup conflict among humans. *Science* **328**, 1408–1411. (doi:10.1126/science.1189047)
16. De Dreu CKW, Shalvi S, Greer LL, Van Kleef GA, Handgraaf MJJ. 2012 Oxytocin motivates non-cooperation in intergroup conflict to protect vulnerable in-group members. *PLoS ONE* **7**, e46751. (doi:10.1371/journal.pone.0046751)
17. Bartz JA, Zaki J, Bolger N, Ochsner KN. 2011 Social effects of oxytocin in humans: context and person matter. *Trends Cogn. Sci.* **15**, 301–309. (doi:10.1016/j.tics.2011.05.002)
18. Kéri S, Kiss I. 2011 Oxytocin response in a trust game and habituation of arousal. *Physiol. Behav.* **102**, 221–224. (doi:10.1016/j.physbeh.2010.11.011)
19. Zak PJ, Kurzban R, Matzner WT. 2005 Oxytocin is associated with human trustworthiness. *Horm. Behav.* **48**, 522–527. (doi:10.1016/j.yhbeh.2005.07.009)
20. Zhong S, Monakhov M, Mok HP, Tong T, San Lai P, Chew SH, Ebstein RB. 2012 U-shaped relation between plasma oxytocin levels and behavior in the trust game. *PLoS ONE* **7**, e51095. (doi:10.1371/journal.pone.0051095)
21. Levy J, Goldstein A, Influs M, Masalha S, Zagoory-Sharon O, Feldman R. 2016 Adolescents growing up amidst intractable conflict attenuate brain response to pain of outgroup. *Proc. Natl Acad. Sci. USA* **113**, 13696. (doi:10.1073/pnas.1612903113)
22. Feldman R, Gordon I, Zagoory-Sharon O. 2011 Maternal and paternal plasma, salivary, and urinary oxytocin and parent–infant synchrony: considering stress and affiliation components of human bonding. *Dev. Sci.* **14**, 752–761. (doi:10.1111/j.1467-7687.2010.01021.x)
23. De Dreu CKW. 2012 Oxytocin modulates cooperation within and competition between groups: an integrative review and research agenda. *Horm. Behav.* **61**, 419–428. (doi:10.1016/j.yhbeh.2011.12.009)
24. Samuni L, Preis A, Mundry R, Deschner T, Crockford C, Wittig RM. 2017 Oxytocin reactivity during intergroup conflict in wild chimpanzees. *Proc. Natl Acad. Sci. USA* **114**, 268–273. (doi:10.1073/pnas.1616812114)
25. Domes G, Heinrichs M, Michel A, Berger C, Herpertz SC. 2007 Oxytocin improves 'mind-reading' in humans. *Biol. Psychiatry* **61**, 731–733. (doi:10.1016/j.biopsych.2006.07.015)
26. De Wilde TR, Ten Velden FS, De Dreu CK. 2017 The neuropeptide oxytocin enhances information sharing and group decision making quality. *Sci. Rep.* **7**, 40622. (doi:10.1038/srep40622)
27. Kohls G, Chevallier C, Troiani V, Schultz RT. 2012 Social 'wanting' dysfunction in autism: neurobiological underpinnings and treatment implications. *J. Neurodev. Disord.* **4**, 10. (doi:10.1186/1866-1955-4-10)
28. Tomasello M, Rakoczy H. 2003 What makes human cognition unique? From individual to shared to collective intentionality. *Mind Lang.* **18**, 121–147. (doi:10.1111/1468-0017.00217)
29. Kaufmann L, Clément F. 2014 Wired for society: cognizing pathways to society and culture. *Topoi* **33**, 459–475. (doi:10.1007/s11245-014-9236-9)
30. Lonsdale AJ, North AC. 2009 Musical taste and ingroup favouritism. *Group Process. Intergroup Relat.* **12**, 319–327. (doi:10.1177/1368430209102842)
31. Beaupré MG, Hess U. 2003 In my mind, we all smile: a case of in-group favoritism. *J. Exp. Soc. Psychol.* **39**, 371–377. (doi:10.1016/S0022-1031(03)00012-X)
32. Kinket B, Verkuyten M. 1999 Intergroup evaluations and social context: a multilevel approach. *Eur. J. Soc. Psychol.* **29**, 219–237. (doi:10.1002/(SICI)1099-0992(199903/05)29:2/3<219::AID-EJSP925>3.0.CO;2-X)
33. Petersen LE, Blank H. 2003 Ingroup bias in the minimal group paradigm shown by three-person groups with high or low state self-esteem. *Eur. J. Soc. Psychol.* **33**, 149–162. (doi:10.1002/ejsp.137)
34. Reichl AJ. 1997 Ingroup favouritism and outgroup favouritism in low status minimal groups: differential responses to status-related and status-unrelated measures. *Eur. J. Soc. Psychol.* **27**, 617–633. (doi:10.1002/(SICI)1099-0992(199711/12)27:6<617::AID-EJSP829>3.0.CO;2-T)