SUPPLEMENTARY INFORMATION FOR: Endogenous oxytocin predicts helping and conversation as a function of group membership

Jennifer Susan McClung*, Zegni Triki*, Fabrice Clément, Adrian Bangerter, and Redouan Bshary

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

†**Correspondence to:** Jennifer McClung

* authors contributed equally

1. Experimental Information

Food Preferences Questionnaire

All items required a response on a 1-7 Likert scale, from "strongly disagree" to "strongly agree".

1. In general, I prefer to eat a cold salad rather than a warm salad.
2. When it comes to sweets, I prefer desserts made with fruit to desserts made with chocolate.
3. I prefer vegetables to fruit.
4. Vegetable soups taste better without cream.
5. I find salty foods can better satisfy my appetite than other types of food.
6. When it comes to dairy products I can never get enough.
7. A meal that consists only of vegetables never satisfies me.
8. In general I prefer savoury to sweet foods.
9. Sometimes I like bitter foods.
10. Meat is a necessary part of my diet.

Experimental Timeline

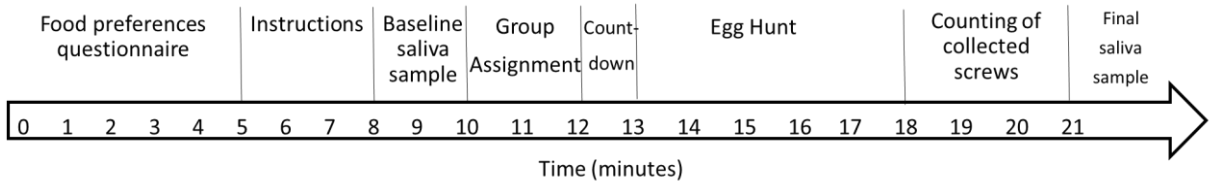

Figure. S1. Schematic timeline of the stages of the experiment.

Experimental Set-up

Figure. S2. A snap shot of the experimental set-up in the lab, including 9 lab tables and 1 window-sill surface, which each had 10 eggs (with 2 blue screws, 2 red screws, and 6 green screws). On tables the eggs were out of players' direct sight (e.g. under objects placed on surfaces or under the tables placed on chairs) whereas on the window sill eggs were visible.

Instructions presented to players

(translated from the French which and presented automatically via a PowerPoint presentation)

1. Welcome

Hello, thank you for coming.

This study is led by Dr. Jennifer McClung of the Biology Department. We want to analyse how participants look for food in a limited physical space, while maximizing their hunting success under time constraints. We will also study the effect that hormone levels can have on these skills. To this end, we will ask you to provide two samples of saliva to conduct the hormonal analyses – one sample now and one at the end of the experiment.

2. Saliva sample 1

3. Food preferences

When you arrived, we asked you to complete a short questionnaire to determine your food preferences. For the purpose of this experiment, we have categorized you either in the apple group or in the orange group. We will be able to compare the performance between the two groups to determine whether there are differences in foraging. The experimenter will now inform you which category you belong to. You will also be asked to wear a lab coat according to your group membership to help us later when we code the videos.

4. Distribution of the lab coats

5. The task

We can now start the task, which simply consists of finding screws that are inside yellow plastic eggs. The screws represent imaginary food. The hunt will take place in this 'field' which is this room. The cameras you are wearing will film your actions throughout the task. This way, we can analyse your use of the space and compare your performance according to your group. Your goal in this experiment is to maximize your profit, i.e. the money you earn at the end of the game.

6. Instructions (1)

Your main task is to search for eggs and collect the screws they contain. The screws are hidden in the little yellow eggs like the one you see next to the computer. The screws must be

screwed into a wooden board that we have been distributed to you for transport – we ask you to carry the screws only after having screwed them to the board to avoid losing them. While hunting you will find screws of three different colours, blue, red and green. One of you will be rewarded for all the red screws collected, the other for all the blue screws collected. The green screws will not be rewarded.

7. Instructions (2)

So even though you are allowed to do what you want with all screws (for example, you could screw all colours into the board) remember that you will be rewarded based on the number of screws of your colour that are screwed into a board at the end of the experiment . Please close and immediately replace the eggs in their original position after inspection/use so that the eggs can then be placed in the same place for the next experiment. (Non-talking condition: You must also complete the task in silence, i.e. without speaking to each other.)

8. Your reward

At the end of the task, you will get CHF 1.- for each screw collected of the colour assigned to you. Since you only have 5 minutes to complete the task, it will be very difficult to fill your board with all the screws of your colour. In order to maximize your success in this task, you are free to organise yourselves as you like, by moving, talking etc. as you like – anything that might help you collect the most screws possible.

As long as you follow these rules, there are no other constraints.

9. Summary

- You must look for blue or red screws but no one is rewarded for green screws
- The screws are rewarded at CHF 1.- each
- Eggs must be closed and replaced in their original place
- Apart from these rules, you are free to complete the task as it suits you

10. Assigning colours

The experimenter will now tell you what colour you are assigned.

The experiment begins at the end of the countdown.

2. Coding Information

Behaviour Coding Scheme

The videos obtained from the eye-tracking glasses were coded using Begaze analysis software (also provided by SensoMotoric Instruments). The videos were used to code players' behaviour and to analyse their conversations (in the talking conditions).

The amount of screws collected was recorded and players' behaviour toward each screw was categorised as either Active helping, Passive helping, or Neglect, as detailed below. Each screw of the other's colour that players found was put into one of these three categories depending on their final decision about what to do with it. This meant that if a player left the other's screw out next to the egg (passive helping) but later went back to collect it (active helping) that screw was categorised as active helping.

1) 'Costly helping' consisted of any action that cost them in terms of time they could spend on their own hunt. This included a) handing the screw to the other player, b) screwing the other's screw into their own board, and c) carrying it with them as they continued the hunt.

2) 'No-cost helping' consisted of any action that helped the other player without costing the hunter any time away from their own hunt. This included a) leaving the egg open with the other's screw visible and b) leaving the egg closed with the other's screw next to the egg itself.

3) 'Neglect' occurred when players did nothing whatsoever but closed the egg again with the other's colour inside it.

4) 'Number of screws collected' was simply the number of screws each player collected of the other's colour. This included screws carried by hand or in their own boards and given to the other player.

Conversation Coding Scheme

All players' conversations from the 1-minute countdown and 5-minute Egg Hunt were transcribed. Talk was segmented into utterances for analysis. An utterance was defined as a unit of speech corresponding to the accomplishment of a speech act or dialogue move (Traum & Heeman, 1997). Utterances typically involve explicit production of a noun-verb-object construction. However, some utterances are elliptical, while others (e.g., back-channel utterances like "uh-huh" or agreement tokens like "ok") may consist of a single word. Each utterance was coded by the first author (who was blind to experimental condition when coding transcripts) into one of the following categories: 1) Shared Intentionality talk, 2) Individual Goal talk, 3) Task talk, and 4) Other talk.

1) Shared Intentionality talk: Shared intentionality involves sharing both a goal and the intentions required to achieve it (Tomasello & Carpenter, 2007). It involves three key components, namely that two people share the same goal, that they both have the intention of completing the shared goal, and most importantly, that they are both aware of their mutual knowledge (or 'common ground') in relation to the goal and the intentions behind it (Manson, Bryant, Gervais, & Kline, 2013; Tomasello et al., 2005). Accordingly, we coded every instance in which players discuss the hunt in terms of shared goals as 'Shared Intentionality talk'. So, any statement in which players discussed the egg-hunt as a shared task with joint goals which depended on the combined efforts of both players fell into this category. An exhaustive list of all types of Shared Intentionality talk players used is as follows:

Planning shared goals prior to hunt
1. ‘We should/could collaborate’
2. ‘We should/could cooperate’
3. ‘We should/could work as a team’
4. ‘We should/could collect (both colours)’
5. ‘We should/could trade screws’
6. ‘We should/could exchange screws’
7. ‘I’ll get yours (yours) if you get (mine)’
8. ‘I’ll (save/keep/bring) yours and you (save/keep/bring) mine’
9. ‘How do you want to do it/organise ourselves?
10. ‘You help me, I’ll help you’
11. ‘We can see what each other opens and benefit from that’

Responses to Shared goal plans during hunt
1. ‘I’ve taken/collected/here’s one (other’s colour) screw/one for you’
2. ‘Let’s exchange/trade/give back the screws’

Collaborative planning about dividing up the room prior to hunt
1. ‘Let’s/we should split the room up’
2. ‘We could/should each check one area’
3. ‘If you do/check (one location) I’ll do/check (other location’
4. ‘Let’s join up/meet at the end’

Reponses to Collaborative plans about dividing up the room during hunt
1. ‘I have/have not checked/done (location in room)’
2. ‘Have you checked (location in room)?’

2) Individual Goal talk: Individual Goal talk included any statement in which a player referenced their separate goals. Individual Goal talk did not address any sharing in goals but referenced the individual’s distinct goals within the hunt (i.e. each individual finding and collecting their own colour). The complete list of all types of Individual Goal talk players used is as follows:

Planning for separate individuals’ goals prior to hunt
1. ‘I’ll tell you where your screws are…’
2. ‘I have to find/collect all (own colour) and you have to find/collect all (other’s colour)’
3. ‘I have (a number) of (my/own colour) screws’

Responses to separate individual goal plans during hunt
1. ‘There is a (other’s colour) screw there’

3) Task talk: Task talk involved any statement about the task itself which did not refer to either individual or shared goals. The list of all types of Task talk players used is as follows:

Practical Statements about the Task / Rules
1. ‘This countdown is ____’
2. ‘This hunt/experiment is___’

3. ‘This is taking (too) long’
4. ‘What is my/your colour?’
5. ‘how do/do we have to screw screws in?’
6. ‘how much time (do we have/is left)?’
7. Egg statements E.g. ‘can we leave the eggs open?’ or ‘these are hard to open’
8. ‘can we start?’
9. ‘do we have to carry the board?’
10. ‘Can I___(other’s colour) screws?’
11. ‘do our groups/colours matter?’
12. ‘Do green screws count?’
13. ‘Do we have the right to/can we talk’

Location of green screws during hunt
1. ‘I’ve found a green screw’
2. ‘There’s a green screw here/there’
3. ‘Green’ upon opening an egg
4. ‘there are empty (eggs)’

4) Other talk: Lastly, other talk involved any other utterances that did not fit into any of the above categories. The list of all types of Other talk is as follows:

Personal Statements
1. ‘Where are you from?’
2. ‘I am from (location)’
3. ‘What’s your name?’
4. ‘My name is…’
5. ‘Where is your accent from?’
6. ‘I_____kinder eggs’
7. ‘___is/am stressed’
8. ‘calm down’

Fillers
1. ‘Yes’
2. ‘Ok’
3. ‘Excuse me’
4. ‘Thanks’
5. ‘Oh crap’
6. ‘perfect’
7. ‘cool’
8. ‘That’s fine’

A second coder naïve to the experimental condition coded 20% of transcripts. To gauge inter-rater agreement we calculated Cohen’s kappa for each of the 4 categories. A kappa of 1 indicates perfect agreement with any kappa greater than 0.7 being considered acceptable. Agreement was high in each category: Shared Intentionality talk $\kappa = 0.83$, Individual Goal talk $\kappa = 0.79$, Task Talk $\kappa = 0.88$, Other talk $\kappa = 0.91$.

3. Hormonal Analyses Procedure:

Salivary Oxytocin Analysis

The Salivettes® absorb non stimulated saliva passively in the participant's mouth in order to have a minimum of 1mL of saliva. Samples were stored immediately on ice until the end of each trial. Salivary OT was analyzed in the lab at the University of Neuchâtel using the enzyme immunoassay technique ELISA. Frozen samples at -20°C were thawed in 4°C. We tried to handle the sample for most of the following steps in 4°C to insure the stability of the salivary OT. The Salivette® tubes were centrifuged for 10 min at 4°C in 5000 rpm. 1 mL of saliva was mixed with 100 µL of phosphoric acid, then centrifuged at 4°C for 10 min at 17000 rpm to avoid pipetting precipitates. At this step, samples can be stored again at -80 °C for further analyses or processed directly for extraction. For ELISA results to be of any use in detecting OT as opposed to other substances in our samples, extraction is crucial (McCullough, Churchland, & Mendez, 2013; Nave, Camerer, & McCullough, 2015). We therefore conducted extraction with an improved extraction protocol employing WCX 25mg 1 mL columns (Evolute SPE, nr. 602-0002-A; for a detailed protocol please see Supplemental Information) where the protocol was tested on pilot samples prior to the experimental samples. In order to obtain the OT levels in the extracted saliva, we used the Enzo Life Sciences® Oxytocin ELISA kit. Kit preparation steps were from the protocol provided with the kit. To improve the results, we used a BioTek EL x 50 microplate washer. For the plate reading we used a BioTek SynergyHT microplate reader. Finally, to analyze the readings we used the software KC4 (V 3.0). Samples were run in duplicate horizontally. The OT concentrations were then estimated from readings of optical density of the standardized curve with serial dilutions of known concentrations. Following instructions provided by the manual in the kit, optical density readings with coefficients of variation (CV) higher than 15 % were excluded from the data set. Furthermore, samples that fell below the detection range were reported as the lowest level of the sensitivity of the kit, 15 pg mL-1 (personal communication with the technical support of Enzo®) only when the lowest limit detection of the plate is equal to or above the sensitivity of the kit. In total 31 samples fell below the lowest range and only 14 were substituted by the value 15pg mL-1 while 16 were recorded as not available. Finally, all OT concentrations were corrected by the concentration factor during extraction. The range of CV% of all the OT readings was between 0 and 8.03%. The lowest detectable level was 0.74 pg mL-1, the intra- and inter-assay coefficient of variations were 2.66 % and 8.33 % respectively.

Solid phase extraction

We collected 0.3-2mL of saliva using Sarsted Salivettes®. Salivettes® tubes were centrifuged at 5000 rpm at 4°C for 10 minutes or more. Then, we mix each 1 ml of the saliva with 100 ul of 0.5 N phosphoric acid in Ependorf® tubes 2 ml or 1.5 ml. Ependorf® tubes were centrifuged in the microcentrifuger at 4°C, 17000 rpm for 10 min and the supernatant was recovered in new Ependorf® tubes. Samples were stored at -80°C for further analyses or processed immediately.

We used Biotage® Pressure +48, positive pressure manifold for solid phase extraction employing mixed mode of non-polar/weak cation exchange columns (Evolute SPE biotage 25mg WCX 1ml, nr. 602-0002-A). At this point we tested the C18 columns as described in the kit protocol but there were no OT detected later in the ELISA plate.

Following the below steps we realized a successful solid phase extraction for the OT peptide:
1) Column activation: condition column with 2ml MeOH (each time 1 ml)

2) condition column with 2ml H_2O pure (each time 1 ml)
3) Loading samples: load sample up to 1 ml, load slowly!
4) wash with 1ml of 95% H_2O + 5% (of a solution of 25% NH_4)
5) wash with 1ml of 95% H_2O + 5% (of a solution of 25% NH_4)
6) Elute slowly with 1.75 ml MeOH 100% for capture of Oxytocin (if using a multi-dispenser, elute with 1ml first followed by 0.75 ml).
7) Speedvac at 45°C at 14 bar for 2 hours, check until dryness
8) store the extract at - 80 °C for further ELISA analyses

Assay procedure

1) The standards: ST1: 900 ul Assay buffer, ST2 to ST7: 500 ul Assay buffer
2) add 100 ul OCT standard in ST1, vortex for ~1 min then take 500 ul from ST1 to ST2.. repeat the same until ST7 (which will be 1 ml volume after the dilutions)
3) reconstitute the extract with 250 ul of assay buffer, vortex well: approximately for 1 min

General advice

a) vortex prior to each pipetting
b) while pipetting try to put it in vertical way without touching the well, avoid the walls as the reagents stay sometimes there.
c) wash the tips in the solution you will pipette before pipetting to have a stable amount of the solution
The steps here are from the provided manual by the kit manufacturer: the plate layout sheet (example Table S1) is prepared prior to each assay as a reference to know where to pipette the designated solutions.
1) vortex the tube prior to each pipetting then Pipet 100 µL of Assay Buffer into the NSB and the Bo wells.
2) vortex the tube prior to each pipetting then pipet 100 µL of Standards #1 through #7 into the appropriate wells.
3) vortex the tube prior to each pipetting then Pipet 100 µL of the Samples into the appropriate wells
4) vortex the tube prior to each pipetting then Pipet 50 µL of Assay Buffer into the NSB wells.
5) Pipet 50 µL of the blue Conjugate into each well, except the Blank wells.
6) Pipet 50 µL of the yellow Antibody into each well, except the Blank and NSB wells.
NOTE: Every well used should be Green in color except the NSB wells which should be Blue. The Blank wells are empty at this point and have no color.
7) Shake the plate gently for 15 min in dark (put it back in the sealing bag).
8) Seal the plate and incubate at 4°C for 18-24 hours.
24 Hours max incubation
9) Empty the contents of the wells and wash 400 µL X 3 times. (test first the washing on pilot plate and clean the filling and aspirating tubes if necessary).
10) Add 5 µL of the blue Conjugate to the TA wells.
11) Add 200 µL of the pNpp Substrate solution to every well.
Incubate at room temperature for 1 hour without shaking.
12) Add 50 µL of Stop Solution to every well. This stops the reaction and the plate should be read immediately.
13) Read the optical density at 405 nm, preferably with correction between 570 and 590 nm. Blank the empty blank and the NSB wells for reading. They will be subtracted from the optical densities values as a control of the machine reading performances.

14) Use the software for the analyses: use 4 parameters curve for values calculation (Fig. S3) as the software will estimate automatically the OT concentrations from the standard curve, the optical density OD is an indication of the optical absorbance: more there are OT in the well lower is the OD reading.
15) Correct the final concentrations by the dilution factor. For example: if you elute 1 ml of saliva and you reconstitute by 250 ul assay buffer, then divide the OCT concentration by 4 to have the real oxytocin concentration in the 1mL of the saliva sample.

Table S1. Plate layout sheet to give an example of the way we prepared the plates for ELISA analysis

	1	2	3	4	5	6	7	8	9	10	11	12
A	BLANK	STD1: 1000.0	STD5: 62.500	SPL2	SPL6	SPL10	SPL14	SPL18	SPL22	SPL26	SPL30	SPL33
B	BLANK	STD1: 1000.0	STD5: 62.500	SPL2	SPL6	SPL10	SPL14	SPL18	SPL22	SPL26	SPL30	SPL33
C	TA	STD2: 500.00	STD6: 31.200	SPL3	SPL7	SPL11	SPL15	SPL19	SPL23	SPL27	SPL31	SPL34
D	TA	STD2: 500.00	STD6: 31.200	SPL3	SPL7	SPL11	SPL15	SPL19	SPL23	SPL27	SPL31	SPL34
E	BLANK (NSB)	STD3: 250.00	STD7: 15.600	SPL4	SPL8	SPL12	SPL16	SPL20	SPL24	SPL28	SPL32	SPL35
F	BLANK (NSB)	STD3: 250.00	STD7: 15.600	SPL4	SPL8	SPL12	SPL16	SPL20	SPL24	SPL28	SPL32	SPL35
G	B0	STD4: 125.00	SPL1	SPL5	SPL9	SPL13	SPL17	SPL21	SPL25	SPL29	SPL33	SPL36
H	B0	STD4: 125.00	SPL1	SPL5	SPL9	SPL13	SPL17	SPL21	SPL25	SPL29	SPL33	SPL36

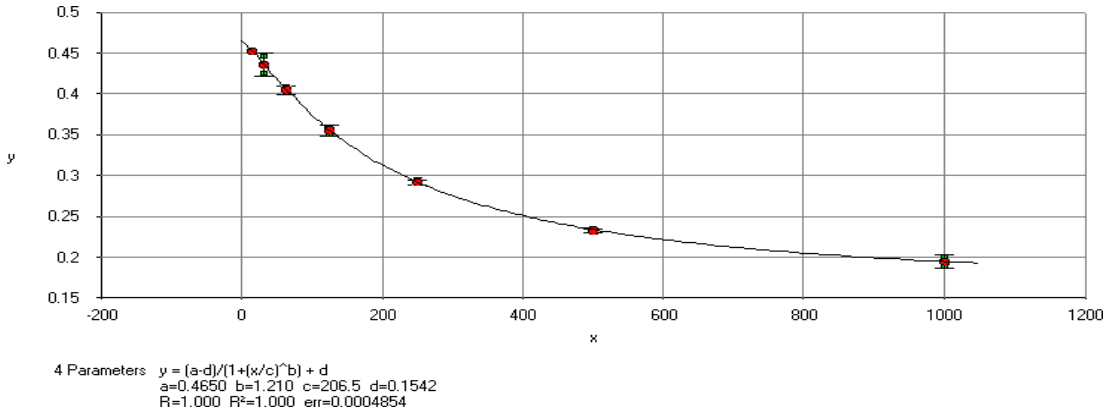

Figure S3. The four parameters standard curve, here is one curve from one plate reading as a demonstration, for each plate we run a separate standard curve as indicated in the manual provided by the kit. Where x is the known concentrations of the serial dilutions in pg/mL, while y is the optical density of the wells. Standards are run here in duplicate.

4. Statistical Analyses

Salivary OT concentrations were measured from the saliva samples collected from salivettes to obtain initial and final levels of OT. This resulted in a total of 219 samples. Due to missing samples and readings that did not fulfil quality requirements, the final sample size of the OT concentrations in the present study was (*n*=61) initial OT concentrations and (*n*=68) final OT concentrations. We obtained paired samples (both initial and final from the same individual) from (*n*=47) individuals in the experiment, and it is this set of samples we used for all analyses.

All data analyses and graphs were generated using the free software R version 3.5.0 (2018-04-23). Data were analyzed with respect to the hypothesis in order to test: (1) whether baseline OT levels could predict cooperative or conversational behaviour during the Egg Hunt, and (2) whether cooperative or conversational behaviour during the Egg Hunt would predict the change in OT over the experiment from baseline to final levels. The experimental design is based on paired players. However, due to missing samples per pair, it was not feasible to run mixed effect models. Bootstrap models with up to 1000 simulations were run instead to check for the robustness of the non-mixed effect models' outcomes. In order to meet the assumptions of normality of distribution and homogeneity of variance of the residuals, arcsine square root transformations were applied to proportional data. We also checked for over-dispersion in the case of the binomial model. Outcomes are reported from the *Anova* (R package "*car*") table with type II analysis of variance (table 1). All continuous predictors such as OT levels and conversation types proportions were standardized to have 0 as a mean using the R function *scale()*. Prior to data analyses we checked for potential bias due to existing differences in baseline OT between genders and/or group membership. The talking condition was not included in the models as a predictor due to missing data (see above). A one-way ANOVA with gender as an explanatory variable revealed no significant difference in baseline OT levels as a function of gender ($F_{(1, 45)} = 0.87$, $p=0.356$). Similarly, no differences in the baseline OT between the in-group and out-group conditions were detected ($F_{(1, 45)} = 1.0647, p=0.308$).

Behavioural analysis

To analyse the relationship between OT and behaviour we used two models: (i) first, a generalized linear model (GLM) with costly helping as the response variable to assess whether baseline levels of OT affected costly helping during the hunt as a function of group membership, and/or gender. The talking condition was not included in the first model as a predictor due to missing data (see above). (ii) second, a Generalized Least Squares (GLS) model with change in OT as the response variable to determine whether cooperation during the hunt would impact the change in OT from baseline to the end of the experiment as a function of group and/or gender. For consistency, the talk condition was not included in the GLS model. Although, we tested for potential significant interaction of the talk condition variable with the rest of the predictors in the GLS model prior to excluding it from the analysis. The GLS model was followed by planned comparisons as post hoc analyses by employing the formula *contrast ()* (R package "*contrast*") with "Holm" adjust method for the *p*-values.

As mentioned in the results section, we obtained only initial or only final samples from some participants (n= 14). These participants were only included in this analysis of baseline OT on cooperation but, because we could not obtain matched samples for these participants, they were missing from subsequent analyses on the change in OT. To test for a potential effect of this missing group, we ran the same analysis on the baseline OT data with two sets of data, all participants including those for whom we did not have matched (before and after) samples (Fig S4) and only those participants for whom we had matched samples (Fig. 1a). This analysis showed that this group of participants did not affect our previous findings as we obtained the same effect ($X^2 = 5.319$, $p = 0.0210$, Fig S4).

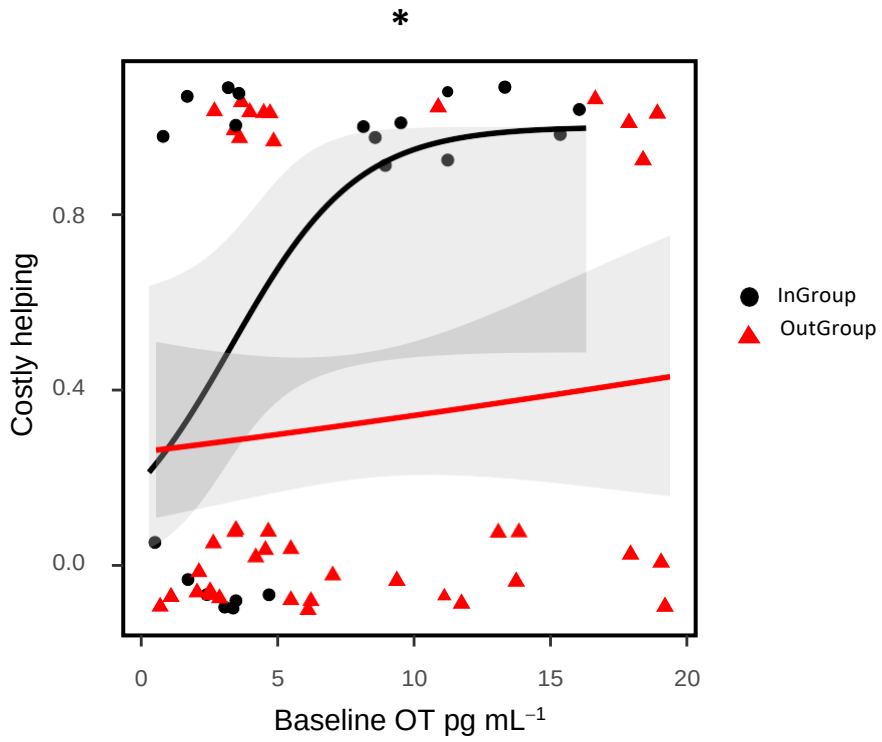

Figure S4. Baseline OT levels predicting costly helping as a function of group membership. Logistic regression lines shown for each condition. Data points are jittered to avoid

overlapping, and the grey area indicates the 95% confidence interval. * $p \leq 0.05$.

Conversation analysis

To analyse the relationship between OT and conversation, we ran a set of linear models (LMs): first, LMs where the conversation types were the response variables to test whether baseline OT would predict each conversation type separately as a function of group membership and/or gender. As above, we also compared the data set in which we had only matched OT samples with the data set of all initial OT samples to assess the correlation between initial OT and types of conversation. The analysis on only participants we had matched samples for is presented in the manuscript, and here we present the analysis on all participants which mirrors the results presented in the manuscript (Shared Intentionality talk: LM: $F(1, 31) = 0.263$, $p = .612$, Fig. S5a; Individual Goal talk: LM: $F(1, 20) = 5.815$, $p = 0.022$, Fig. S5b; Task talk: LM: $F(1, 20) = 0.417$, $p = .523$, Fig. S5c; and Other talk: LM: $F(1, 20) = 2.460$, $p = .126$, Fig. S5d).

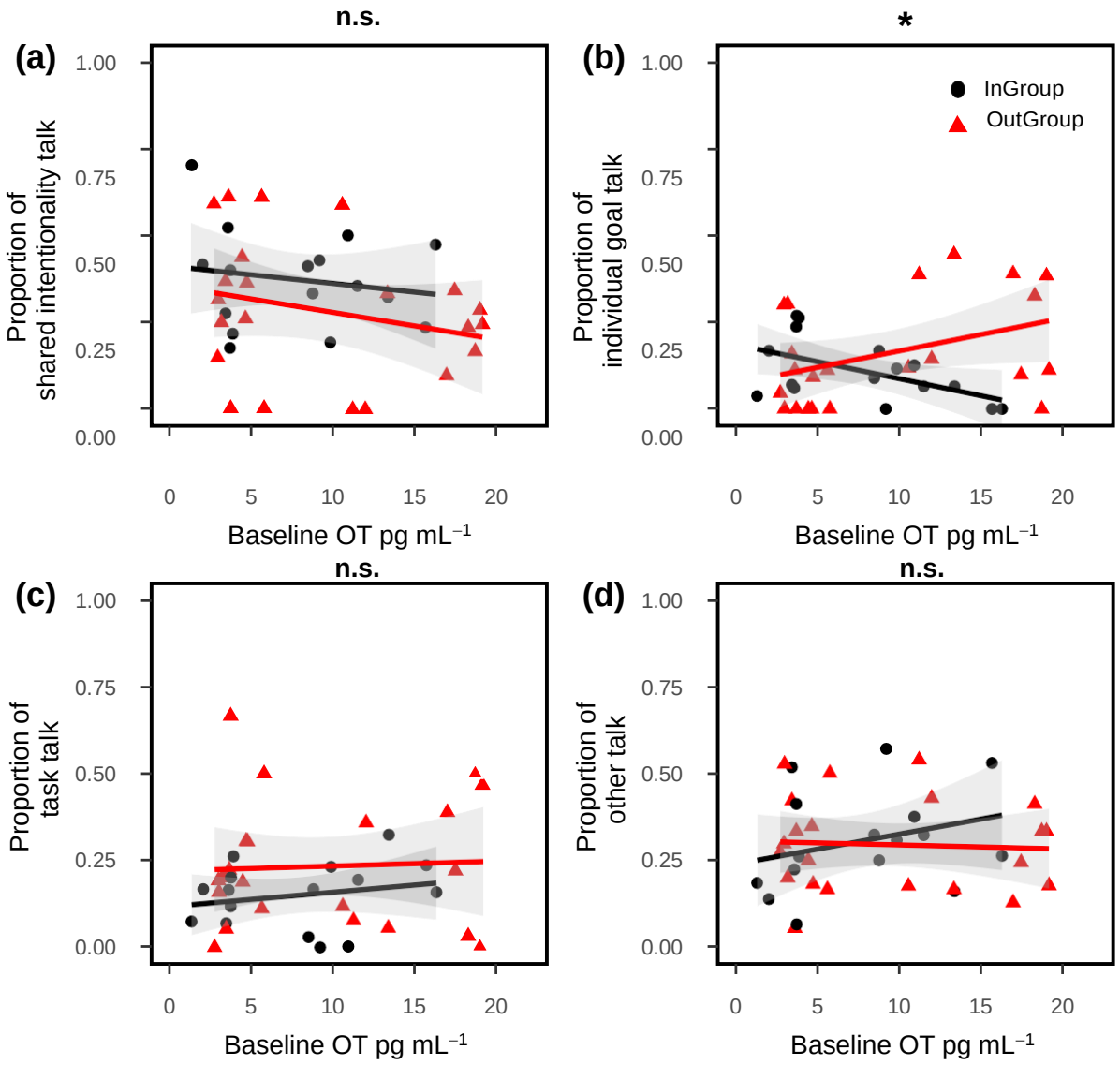

Figure S5. Baseline OT levels predicting different types of talk as a function of group membership with 95% confidence intervals, analysis on all participants including those for whom matched samples were not obtained. (a) Shared Intentionality talk; (b) Individual Goal talk; (c) Task talk; and (d) Other talk.
Logistic regression lines shown for each condition. ** $p \leq 0.01$; n.s. = non-significant differences.

Lastly, to explore the relationship between individual goal talk and costly helping, we ran a GLM and found a negative correlation (GLM: X^2 = 6.026, p = 0.014, pseudo r^2 = 18%; Fig. S5).

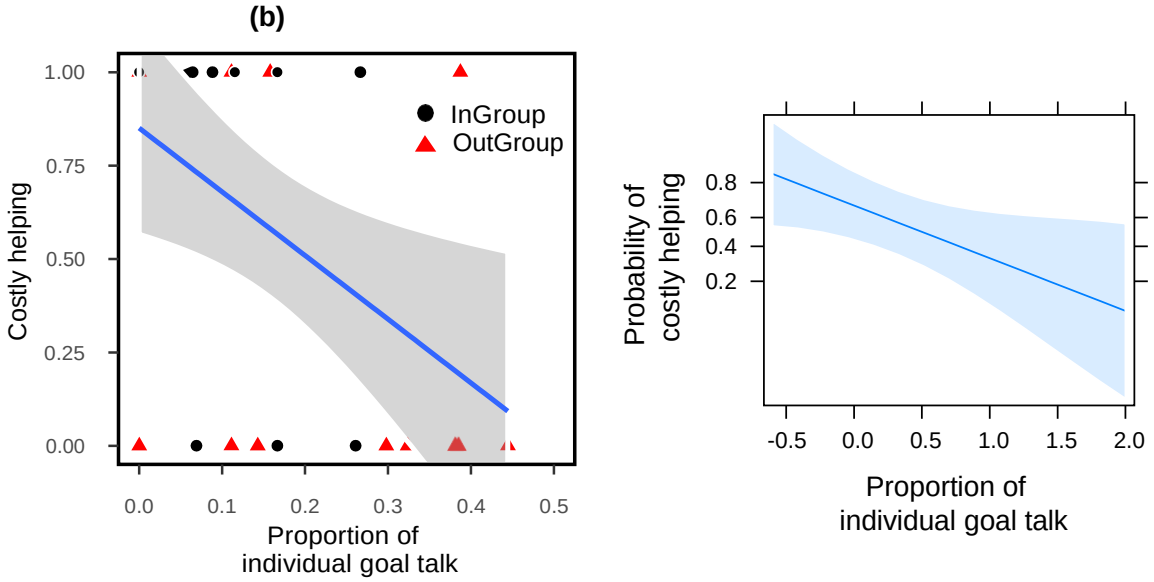

Figure S6. Predictions of costly helping predicted by individual goal talk. (a) Scatterplot of the raw data. (b) Plot effect showing the predictions by the model.