**Oxytocin, individual differences, and trust game behavior: a registered large-scale replication**

Charlotte F. Kroll1,2,3, Koen Schruers1, Wolfgang Viechtbauer1, Claudia Vingerhoets1, Leonie Seidel4, Arno Riedl5, and Dennis Hernaus1

1Department of Psychiatry & Neuropsychology, School for Mental Health and Neuroscience, Faculty of Health, Medicine and Life Sciences, Maastricht University, Minderbroedersberg 4-6, P.O. Box 616, Maastricht 6200 MD, The Netherlands

2Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, 6200 MD, Maastricht, The Netherlands

3 Department of Microeconomics and Public Economics (MPE), P.O. Box 616, Maastricht 6200 MD, The Netherlands

4Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

5CESifo, IZA, Netspar and Maastricht University, Department of Microeconomics and Public Economics (MPE), P.O. Box 616, Maastricht 6200 MD, The Netherlands

\* corresponding author:

Charlotte F. Kroll

charlotte.kroll@maastrichtuniversity.nl

**Abstract**

The neuropeptide oxytocin (OXT) is thought to modulate important aspects of prosocial behavior. In a seminal paper, Kosfeld et al. (2005) reported that intranasally administered OXT modulated trusting behavior in an economic trust game. Several attempts to conceptually replicate these findings yielded mixed results, which might be partly due to small sample sizes that can reduce the ability to detect, or reject, meaningful effects. Here, we propose to perform a large-scale replication (*N*=220) of Kosfeld et al. (2005) with specific attention for small effects and subpopulations whose trusting behavior may be sensitive to OXT manipulations. Moreover, we will conduct the largest-ever pooled analysis by merging our data with data from a previous replication by Declerck et al. (2020). Using additional (equivalence) analyses, we aim to refute effect sizes of OXT on interpersonal trust that will not be worthwhile pursuing in most lab-based contexts. Our study will contribute to a more refined understanding of OXT’s involvement in human social behavior, for example by identifying boundary conditions that will delineate when OXT-induced effects on prosocial behavior may occur. Critically, we anticipate that our work will offer a more realistic perspective on the effect sizes that can be expected when using intranasal OXT to modulate prosocial behavior.

*Keywords:* Oxytocin,trust game, replication, individual differences, pooled analysis

**Introduction**

The hormone oxytocin (OXT) is a nine amino acid neuropeptide that is mainly synthesized in the hypothalamus and that acts both centrally and peripherally1,2. In both the biological and social sciences, it has been extensively linked to human social behavior3-5. For example, its role in labor6,7 and lactation8,9 is well established, and it is further linked to parent-infant10,11 and romantic12 bonding. More recently, OXT has garnered considerable interest for its purported involvement in complex social cognition and behavior13-15, including empathy16,17, altruism18,19, and emotion recognition20,21, which have all contributed to OXT’s increasing reputation as the ‘prosocial hormone’22. This idea is further fueled by studies that have revealed how administration of intranasal, synthetic, OXT modulates the neural correlates of these complex social behaviors, including interpersonal trust, social value representation, and perspective-taking5,23,24. Moreover, intranasal administration of OXT has been shown to impact neural activation more generally associated with appetitive and/or aversive processes (e.g., salience attribution, representing reward value, and valence) in the brain’s canonical reward network, including ventral striatum, ventromedial prefrontal cortex, and ventral tegmental area23,25-28. Together, these results have sparked a wave of clinical, e.g.,2,29,30, and non-clinical, e.g.,3,4,31, studies that have used intranasally administered OXT to influence various aspects of social behavior32.

One aspect of social behavior that has received particular interest in relation to OXT is interpersonal trust, e.g.,3,4,33, a psychological state that entails the risk to be exploited weighted against expectations of other’s prosocial behavior34. On a behavioral level, an established instrument to measure interpersonal trust is the trust or investment game35,36. This game is typically played in dyads, where one player is assigned the role of investor, and the other the role of trustee. The investor receives an initial endowment and decides how much of it, if any, to send to the trustee. This sent amount is tripled and added to the initial endowment of the trustee, who then decides how much, if any, to send back to the investor. Since the investor cannot be certain about any back-transfers, the process of investing money represents the investor’s willingness to accept the risk of being exploited weighted against expectations of the trustee’s prosocial behavior (i.e., the amount sent back), which is interpreted as a proxy of trusting behavior35.

In a seminal paper from 2005, Kosfeld and colleagues3 reported that intranasally administered OXT modulated the degree to which humans trust each other during the trust game. Specifically, investors who received intranasal OXT sent more money to trustees than investors who received a placebo nose spray. Several attempts to conceptually replicate these findings using different designs (e.g., playing with a real vs. computerized trustee, within- vs. between-subjects), and set-ups (e.g., magnetic resonance imaging vs. behavior only) have led to mixed results, e.g.,23,33,37,38. The reported findings range from a trust increase, to no effect, to a trust decrease following OXT administration39, suggesting that the effect is less robust than originally assumed37,39,40.

Crucially, the original study as well as many of the replication studies that ensued were underpowered39,40. If only studies with small sample sizes that happen to achieve statistical significance are published, which can only happen when the effect size is large, i.e., under publication bias41, biased effect estimates can arise. Both factors (low sample sizes, publication bias) make meta-analyses prone to effect size inflation, which should therefore be interpreted with caution42,43.

In light of the mixed results, well-powered replications are crucial, and should ideally be conducted with methods that closely match those employed in the original studies44, and make materials and results publicly available45. Moreover, sufficiently large sample sizes can help differentiate between true null effects and significant effects that are too small to be considered meaningful (e.g., when using equivalence testing46). For example, Quintana47 used equivalence testing to show that more than 25% of non-significant meta-analytic effects were a consequence of data insensitivity, rather than true equivalence between OXT and placebo groups. These observations fit with the idea that many previous intranasal OXT studies could not reliably detect or reject meaningful effects42.

Recent work has started to address these important issues. Declerck et al.4 conducted a pre-registered, well powered, replication study of Kosfeld et al.3, which aimed to replicate the effects of intranasal OXT on interpersonal trust in a sample of 667 participants4. These participants were randomized to a minimal social contact condition, in which they could talk to their potential game partners before the start of the experiment and that matches the respective condition used in the original study by Kosfeld et al.3 (*n*=326), as well as a no-social-contact group in which participants were not allowed to engage with each other before the trust game. Importantly, as hypothesized, intranasal OXT, compared to a placebo, did not increase interpersonal trust in either of these conditions.

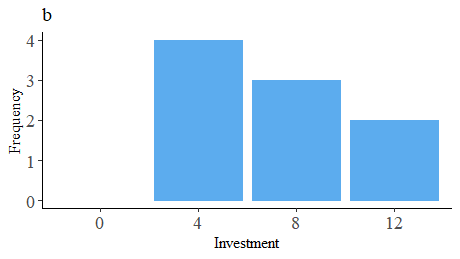
Although the results of this study cast doubt on the notion that intranasal OXT administration has substantial effects on interpersonal trust, there is the possibility that OXT may have a smaller effect, perhaps limited to particular subpopulations. For example, a non-preregistered follow-up analysis by Declerck et al.4 suggested that OXT increased trust in a subsample of participants with a low disposition to trust, measured using a validated self-report questionnaire36. Moreover, effects of intranasal OXT on prosocial behavior, and associated neural correlates, have also been reported to be associated with interindividual (trait) differences in self-regarding vs. other-regarding attitudes (e.g., social value orientation5). Such interindividual differences likely are not restricted to traits that are important in social contexts: in a host of other studies, the effects of intranasal OXT have been reported to be mediated by interindividual variation in more general reward/approach and punishment/avoid tendencies (e.g., behavioral activation and inhibition systems, or sensitivity to rewards and punishments48-51). These observations align with the idea that OXT might also modulate more domain-general reward sensitivity and associative learning capacities across a wide range of paradigms, perhaps reflecting an increase in sensitivity to positive non-social outcomes26,27,52-54. Additionally, these observations might follow from the basic premise of the trust game35, where investors may be tempted to exhibit potential trusting behavior (i.e., investing) because they can earn greater monetary rewards when their trust is reciprocated. Indeed, Mislin and colleagues55 reported that during the investment game, higher possible rewards may increase the likelihood of trusting behavior, fitting well with neurobiological studies that have demonstrated reward-related activation during investment games56,57. Thus, OXT effects on interpersonal trust may be dependent on socially specific and more general individual psychological differences. This is in line with the importance of considering heterogeneity in study populations to establish comprehensive explanations of potential causal mechanisms58.The discussed considerations, most notably the growing skepticism regarding OXT effects in the context of social behavior, e.g.,39, clearly indicate the need for further transparent investigation and replication of intranasal OXT studies using a sufficiently large sample size and applying proper statistical techniques to improve interpretability such as equivalence testing to more reliably detect or reject meaningful effects46 (e.g.,39,40,44). Furthermore, the role of individual trait differences such as baseline trust levels or baseline sensitivity to potential rewards should be additionally considered, e.g.,4,59,60, and expectations regarding their role should ideally be pre-registered in order to limit the possibility that such associations are identified based on exploratory *post hoc* analyses.

We propose to perform a large-scale pre-registered replication study of the seminal study by Kosfeld et al.3, with additional attention for subpopulations whose behavior may be especially sensitive to OXT. In line with previous replication attempts, e.g.,4,39,40 we do not expect to find significant differences between OXT and placebo in overall investments in the trust game (*Hypothesis 1a*). Even if, contrary to the prediction, we would observe a significant difference between the OXT and placebo group, we expect it to be significantly smaller than a predefined effect size of interest46 (*Hypothesis 1b*). Next, regarding the role of individual differences, we expect the effect of OXT on investments to decrease with increasing trust propensity of the participants (*Hypothesis 2*). Further, we expect the effect of OXT on investments to decrease with increasing reward sensitivity and decreasing punishment sensitivity (*Hypothesis 3*), reflecting a trust-enhancing effect of OXT in individuals with lower behavioral activation tendencies or higher behavioral inhibition tendencies. Lastly, we will capitalize on the publicly available data from a previous large-scale replication that used a similar design, by pooling our data with the data from Declerck et al.’s minimal social contact condition61 to increase power to find a small effect of OXT on intrapersonal trust. To achieve this, we will re-run Hypothesis 1 using this combined dataset (we refer to this as “pooled analysis”).

**Methods and Materials**

all 220 (see *Participants* section below for details) within a pre-defined time interval, where half of the participants will be randomized to placebo, and the other half to OXT (between-subjects randomization: treatment; main randomization factor, 50:50 whenever possible). During each session 2, E

experimental money units (EMU)



2

**Participants**

We will recruit 220 participants (*n*=110 randomized to placebo, *n*=110 randomized to OXT) aged 18-32, using subject pools available at Maastricht University and via online study advertisement. All participants will take part in the two sessions described under *Study design*.

A sample size of *N*=220 is considerably larger than the one collected by Kosfeld et al.3, who conducted their key analysis on a sample of 29 vs. 29 investors on/off OXT. Importantly, and in contrast to other OXY-trust studies, we base our sample size on an *a priori* simulation power analysis. To this end, we used Figure 2a from the original Kosfeld et al.3 publication to extract the probabilities of choosing the different investment amounts. Here, two things are noteworthy. First, due to the design specifics of Kosfeld et al.3, the reported values in *Figure 2a* are mean investments over four rounds of the trust game. Second, as noted by Calin-Jagermann and Cumming63, the bar heights for the placebo group in the figure sum up to less participants than reported in the text, which they suspect (and we confirmed by changing the background color of the pdf) to result from a clipped or misprinted aspect of the figure. They imputed three additional scores of a mean investment of 10 in the control group, which yielded a dataset that could reproduce all statistics reported by Kosfeld and colleagues3.

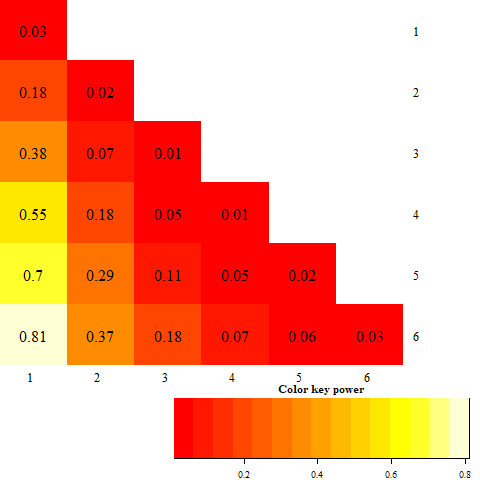
We used the publicly available reconstructed dataset by Calin-Jagermann and Cumming64 for probability extraction. To derive probabilities for the investment amounts of 0, 4, 8, and 12 EMU, we collapsed probabilities in line with Kosfeld et al.3 such that the probability of an investment of 0 consisted of the probabilities of (rounded) mean investments of 0, 1, 2, and 3; 4 consisted of 4, 5, 6, and 7; 8 consisted of 8, 9, 10, and 11; and 12 of only 12 due to its high frequency in the original study3.In light of our main hypothesis (*Hypothesis 1*) – that is, the hypothesis pertaining to monetary investment on/off OXT – we extracted probabilities from both the OXT and placebo condition of Kosfeld et al.3 and simulated 1000 random data sets to derive the empirical power for detecting a significant group difference based on an ordinal logistic regression (OLR) model, fitted using the polr() function from the *MASS*65 package in *R*66. We chose an OLR because our dependent variable (investments) has only four levels (0, 4, 8, and 12) and the (subjectively perceived) difference between those investment values may not be equal across individuals. Using the reported probabilities from Kosfeld et al.3, and with *α* set to .02 (testing one-sided; i.e., greater monetary investments on OXT compared to placebo), a sample of 220 participants would yield a power of .99 to detect an increase in monetary investments (i.e., interpersonal trust) following OXT administration. Applying the packages *brant*67 and *nnet*65, we found an average rejection rate of the proportional odds assumption of .1 in this approach, which we deem acceptably low.

Critically, however, mixed results from previous OXT-trust studies suggest that the true effect regarding an increase in interpersonal trust following OXT administration, e.g., 4,37,39 is smaller than the effect size of Cohen’s *d* of .51 (corresponding to an odds ratio of 2.5268) reported in the original study by Kosfeld et al.3. One reason for a smaller true effect size could lie in the details of the paradigm, which leads to skewed investment distributions. Specifically, approximately 55% of participants in the original study3, and almost 80% in a recent replication study4, invested 8 or the maximum of 12 EMU in the placebo condition, which substantially decreases the power to detect any trust-enhancing (i.e., greater investments) effect of OXT.

If the true effect was indeed smaller than the one reported in Kosfeld3, all studies conducted to date would be severely underpowered. We propose a way to partly compensate for this problem by combining our data with the minimal social condition sample from a previous replication study with publicly available data61 (*n*=326; *Pooled analysis*). We used extracted probabilities for the placebo condition from Kosfeld et al.3 and assumed a more reasonable true minimal effect size, a Cohen’s *d* =.2, which would translate to an odds ratio of 1.4468. We further believe that demonstrating an effect size smaller than *d*=.2 is infeasible in most human lab studies. For instance, a Cohen’s *d* of .1 would require a sample size of around 2500. Next, we again simulated 1000 random datasets to derive power. Using the proposed study design (see below), and with *α* set to .05 (testing one-sided), a combined sample of *N*=546 (220 participants collected as part of this pre-registration) would yield a power of .8 to detect an increase in interpersonal trust following OXT administration. Lowering *α* to .02 (again, one-tailed) would still yield a power of about .68.

As opposed to the power simulation for the main hypothesis, the power simulation for the treatment-by-trait score interaction hypotheses is more challenging, primarily because we do not have reliable priors for data simulation. The only well-powered study to date that examined an interaction between OXT (vs. placebo) and trust propensity on investments is the recent replication by Declerck and colleagues4. Moreover, to our knowledge, there are no well-powered datasets to conduct an informed power analysis for interactions with sensitivity to reward and punishment. We therefore chose to report a power spectrum analysis for Hypotheses 2 and 3 by estimating the power for a range of (treatment-by-trait) interaction effect sizes, in line with previous work from Quintana69. To this end, we simulated the power for a range of differences in odds ratios for the treatment effect (i.e., OXT vs. placebo effect on investments) between people that score high vs. people that score low on the trait.

Similar to the main power analysis, we used the investment probabilities of the placebo condition from Kosfeld et al.3 as a starting point to simulate investment probabilities for people who score low, compared to people who score high on a hypothetical trait that may reflect trust propensity, sensitivity to punishment, or reward (for simplicity, a median split is used to create low versus high trait groups). Next, we simulated 1000 random datasets to derive our power for each unique combination of a difference in the treatment effect(odds ratio of 1=very small, odds ratio of 6=very large) for the low and high score trait groups. With alpha set to .02, we obtain the power spectrum results as presented in *Figure 3*.



**Figure 3.** Power spectrum analysis for interaction hypotheses 2 and 3. Values of 1-6 horizontally and vertically to the matrix depict the odds ratios of the treatment effect (i.e., investments on OXT vs. placebo), one for the low and one for the high trait group. Values within the matrix refer to the power to detect a treatment-by-trait interaction in the model of investments, which thus depends on the magnitude of the difference in the treatment group for the high vs. low trait group. For instance, if there would be a treatment effect of odds ratio=1 in one group, and a treatment effect of odds ratio=5 in the other group, we would obtain a power of .7 to detect a treatment-by-trait interaction on investments.

As expected, this analysis demonstrates that the power to detect a significant interaction depends on the magnitude of the difference of the treatment effect for individuals who score high, compared to individuals who score low, on the trait. With our planned sample size of N=220, only large differences in the odds ratio would yield sufficient power to detect a treatment-by-trait interaction on investments. Importantly, this is not an entirely unrealistic situation: Let us assume that approximately 50% of individuals score low and 50% score high on a trait and take into account that Declerck et al. (2020) reported that the effect of oxytocin on investments diminishes with higher dispositional trust ratings. Consequently, to retrieve the effect reported in Kosfeld et al.3, i.e., an effect size of Cohen’s *d* of .51 which approximately translates to an odds ratio of 2.5268, the discrepancies between odds ratios in the two groups would have to be large.

Strictly in line with Kosfeld et al.3, recruitment will be limited to male participants. Other exclusion criteria in line with Declerck et al.4 will be a lifetime diagnosis of a DSM-V psychiatric (including substance dependence and abuse) or neurological disorder, current treatment by a psychologist or psychiatrist for mental health-related problems, and current use of psychoactive medication for mental health-related problems, that is, antidepressants, antipsychotics, benzodiazepines, anxiety medication, neuroleptics, anticonvulsants, and stimulants. Further, participants who are allergic to latex will be excluded70. Participants will be screened on these criteria via a digital screening form. To minimize side effects associated with OXT-induced urinary retention, a consumption maximum of 1 liter of water or other liquids 2 hours before session 2 will be required. In case participants fail to adhere to this rule, they will be rescheduled for a later session. Participants who fail the water consumption guidelines on two separate occasions or who fail twice to complete session 1 (i.e., after rescheduling), will be excluded from further participation. Those who fail to complete the investment paradigm, report to have severe nose obstruction on the day of the experiment, or did not comply with rules abstaining from alcohol, non-prescription drugs, and smoking (see *Experimental procedure*) will be excluded from data analysis. Recruitment will continue until 220 participants have successfully completed session 1 and 2. If a sufficient amount of participants cannot be recruited in the Maastricht region, further sessions will be run in labs at other universities.

Approval of the study will be obtained from the *Maastricht University/Maastricht Academic Medical Ethics Committee* and participants will provide written informed consent prior to starting session 1. They will be reimbursed with a flat fee of €20 plus their earnings from the trust game, which range from €0 to €72 depending on their behavior, and their earnings from an additional generalized dictator game, which range from €0 to €10.

**Oxytocin administration**

Synthetically produced OXT, administered non-invasively via a nose spray, will be used to temporarily manipulate OXT levels. In previous research, it has been assumed that intranasal administration of OXT crosses the blood-brain barrier via a layer of cells in the nasal cavity epithelium, where it can travel to the hypothalamus1, thereby exerting central nervous system effects. However, a recent review71 suggests that intranasally administered OXT may primarily exert its central effects via the nose-to-brain route, rather than crossing the blood-brain barrier. Even though direct confirmation of this route in humans is yet to be established, evidence suggests that intranasally administered OXT can reach the central compartment72 via diffusion-mediated transport, and can then bind to OXT receptors throughout the brain73.

In line with previous work4,74,75 and the original study3, participants in the OXT treatment group (*n*=110) will receive a commonly-used dose of 24 I/U of *Syntocinon* (brand name) in 6 puffs (three per nostril) or an equivalent number of puffs of a hypotonic saline, placebo, spray with similar ingredients but no OXT. Although previous studies have reported dose-dependent effects of OXT on psychological measures76, consistent with the original study by Kosfeld et al.3, and for feasibility reasons, we use a single dose of 24 I/U here. Administration will be performed strictly in line with the recommendations and guidelines provided in77, including standardized demonstration, test puff, and visual inspection by the experimenter. *Syntocinon* has a relatively short-lived pharmacokinetic profile with concentrations of OXT peaking between 30-90 minutes post-administration; it is subsequently cleared rapidly from the body and not detectable around 150 minutes post-administration31. The trust game will therefore be completed during the OXT peak at 50 minutes post-administration in line with the original study3.

**Experimental procedure**

Our experimental procedure will be partially based on the replication by Declerck and colleagues4, with adjustments made to serve the purpose of our study. Specifically, compared to the replication study by Declerck et al.4, the current replication effort differs in terms of questionnaires (omission of some questionnaires that did not directly address objectives of the current study; addition of questionnaires, most notably the SPSRQ-RC78 for testing our third hypothesis), reimbursement (guaranteed payout of all trust game decisions to increase fairness, as opposed to only guaranteed payout of investor decision in Declerck et al.4), the collection of saliva samples (for intended analysis of hormonal markers), and the addition of a generalized dictator game79 in session 2, of which participants will not be aware when playing the trust game.

Prior to starting session 1, participants will receive a screening form assessing all study exclusion criteria (see *Participants*). Eligible participants will provide written informed consent and afterwards receive a secure *Qualtrics* link to complete session 1. In this session, they will complete a battery of questionnaires assessing traits, states, and other subjective experiences (an overview of all questionnaires used can be found in *Table 1*).

|  |  |
| --- | --- |
| Variable | Questionnaire |
| Dispositional trust | Inclusive Generalised Trust Scale (IGTS36) |
| Sensitivity to punishment and rewards | Sensitivity to Punishment and Reward Questionnaire – Revised and Clarified (SPSRQ-RC78) |
| Chronic nasal obstruction | Nasal Obstruction Symptom scale (NOSE80) |
| Social skills and competencies | Autism Spectrum Quotient 10 scale (AQ1081) |
| Social value orientation | Triple dominance measure for social value orientation (SVO82) |
| Optimism | Life Orientation Test – Revised (LOTR83) |
| Various types of impulsiveness and inattentiveness | Abbreviated Impulsiveness Scale (ABIS84) |

**Table 1.** Self-reported symptoms and traits assessed during session 1. Questionnaires will be presented randomized at the individual level.

Baseline measures of dispositional trust and sensitivity to punishment and rewards will be used when testing our registered hypotheses, with levels of chronic nasal obstruction (NOSE80) as a covariate, whereas the remaining measures will serve as moderators in potential further exploratory analyses.

Participants will have a maximum of two attempts to complete session 1 at least three days before the start of session 2 (i.e., one first attempt, and a second and final attempt in case attempt 1 failed for any reason).

Session 2 will take place in the laboratory and last a maximum of 2 hours. Upon arrival at the lab, participants will first indicate compliance with the water consumption guidelines. In case of positive compliance, they will be randomly allocated to a sight-shielded cubicle with a laptop. When seated they will complete a multidimensional mood state questionnaire (MDBF85) to assess momentary mood states. Next, two saliva samples will be collected, the first one using *Sarstedt Salivette* cotton swabs, the second one using *Oragene Discover OGR-500 kits* (*DNA Genotek*; DNA will be extracted using *prepIT•L2P reagent* (*DNA Genotek*) and quantified with *PicoGreen* (*Quant-iT PicoGreen dsDNA* Assay Kit, *Thermo Fisher Scientific*). Participants then receive standardized verbal and written OXT self-administration instructions (see *Appendix C*) according to77 and the experimenter will demonstrate the use of the nasal spray.

Subsequently, participants self-administer the nose spray using a metered finger spray, under the supervision of trained experimenters, and rate the spray administration discomfort. They are not informed about any potential effects of OXT on prosocial behavior since this would reveal the purpose of this experiment. In order to fill the waiting time until peak OXT concentrations, participants will complete the Extraversion scale of the HEXACO-10086 to assess extraversion and social competencies, the Negative reciprocity subscale of the Global Preferences Scale87 to assess inclination to negative reciprocity, and the Physiological Arousal Questionnaire (PAQ88) to assess momentary arousal (with order randomized at the individual level). After completion of these questionnaires, participants will be directed from their cubicles to a common room and will be instructed that they can now to talk to each other quietly, in order to establish minimal social contact as in the original study3. They will stay together in the common room for 10 minutes. After that, they will be guided back to their cubicles. Here, participants will fill out the MDBF85 once more and provide another saliva sample, collected using *Sarstedt Salivette* cotton swabs. Thereafter Fifty minutes post-administration of OXT, they will make choices in the two previously described decision cycles of the trust game (see *Trust game*).

After finishing the trust game, participants will again complete the MDBF85 and provide a final saliva sample using *Sarstedt Salivette* cotton swabs. Next, participants will receive standardized instructions for the dictator game79 and make choices in two practice dictator game decision problems. They will then complete 50 decision problems of the dictator game (i.e., divide an endowment between themselves and a randomly chosen receiver) to assess social preferences. Payoffs will be determined in line with Fisman et al.79; that is, the experimental program will randomly select one of the 50 decision problems and payments will be according to the allocation made in the selected problem.

At the end of session 2, participants will rate whether they think they have received OXT or placebo. In line with Declerck et al.4, participants will additionally be asked how connected they felt to their trustee using the Inclusion of the Other and the Self (IOS) scale89. Next, they will report whether any friends or acquaintances were among the other participants, whether they have previously participated in a similar experiment, how well they understood the instructions of both games, the extent to which they think the instruction of the games were truthful, and whether they abstained from alcohol, non-prescription drugs, and smoking more than 20 cigarettes 12 hours prior to session 2. Lastly, participants will rate their current nose obstruction on a scale from 0 (incredibly clear) to 10 (incredibly blocked). A score of >8 will be considered as severe nasal obstruction90.They will then fill in a reimbursement form and will be dismissed.

**Statistical analyses**

Statistical analyses will be performed using *R* version 4.2.366. We do not expect missing values for investments since this will be the primary outcome measure, and a failure to collect these data will lead to exclusion from the study. Further, given the nature of the primary outcome variable with only four possible levels (0,4,8,12), there is no possibility for a single value to deviate strongly from the bulk of distribution and hence, there is no need to define and remove outliers.

First, visual inspection of floor and ceiling effects will be performed as quality checks. A floor effect (i.e., no investment made in the placebo condition is frequent) would be unproblematic, since our directional, one-sided, hypotheses predicts an increase, not a decrease, of interpersonal trust after OXT administration. In contrast, a ceiling effect (i.e., the highest investment is chosen frequently in the placebo condition), would decrease the power to detect an effect of OXT on investments. We will report the distribution of investments and formally evaluate the presence of a ceiling effect using the proportional odds assumption65,67, i.e., that the effect of an independent variable is constant for each increase in the level of the response (here, investment levels). In case of the presence of a ceiling effect based on this criterion *(brant*67 and *nnet*65 test reveal a *p* ≤ .05), we will conduct three separate binary logistic (BL) models, comparing the individual investments (I: 0=0 and 4/8/12=1; II: 0/4=0 and 8/12=1; III: 0/4/8=0 and 12=1) between the OXT and the placebo group (see *Hypothesis 1a*). Pilot data collected between January and June 2021 from a total of 16 participants (*n*=9 in the placebo group) did not reveal visual floor or ceiling effects (see *Figure 1b*).

IGTS scores will be calculated by taking the mean of the nine items, measured on a 7-point Likert scale. The reliability of the IGTS has been found to be acceptable (*α*=.83 in a population of 600 non-student residents between 20-59 years with an approximately equal sex ratio for each 10-year age cluster who were living in a wealthy neighborhood in Japan) and the scale correlates moderately with behavioral measures of trustworthiness (*r*=.25) and trust (*r*=.29)36. SPSRQ-RC scores will be measured using 20 items scored on a 5-point Likert scale; odd item scores will be summed to create a sensitivity to punishment (SP) score, even item scores will be summed to create a sensitivity to reward (SR) score. The reliability of both subscales is acceptable (*α*=.86 for the SP subscale, and *α*=.82 for the SR subscale in a population of 796 participants between 18-54 years (71% female) that were enrolled in courses at a large East Coast university in the United States78). NOSE scores will be measured over five items on a 5-point Likert scale (ranging from 0 to 4), summed, and multiplied by 5 to allow for a range of scores between 0 and 100. These scores are used to categorize mild (range 5-25), moderate (range 30-50), severe (range 55-75), or extreme (range 80-100) nasal obstruction80. Missing questionnaire scores will be treated as missing values in regression analyses, and the respective participants will be omitted from analyses that involve the particular questionnaire.

Descriptive statistics will include the number of participants per condition (OXT vs. placebo) as well as the mean and *SD* of the IGTS36 and SPSRQ-RC78 scores, and the frequency of nasal obstruction levels80.

Statistical tests will be considered significant at *p* < .02 for our replication analyses (Hypotheses 1-3), and *p*<.05 for our pooled analysis. For the latter, we deviate from the stricter threshold since we aim to balance specificity and sensitivity for our analysis; maintaining an alpha of .02 would necessitate a sample size that exceeds our data collection capacity. For example, detecting a small effect size of Cohen’s *d*=.2 for our pooled analysis, would require around 850 participants with a power .8 and alpha .02. We will supplement the test statistics with Partial Eta Squared values (*ηp*2) as a measure of effect size (*ηp*2 of .01 indicates small effects, *ηp*2 of .06 medium effects, and *ηp*2 of equal to or greater than .14 large effects). When testing our registered hypothesis, we will not additionally correct for multiple comparisons since we are using a stricter threshold alpha of .02 for all replication-related hypothesis testing. This alpha of .02 is close to a Bonferroni-corrected alpha of .05, corrected for the number of hypotheses. In any exploratory analyses, we will correct for multiple comparisons and report uncorrected and corrected test results.

***Hypothesis 1a***

To investigate if investors on OXT, compared to placebo, do not significantly invest more money in trustees (dependent variable) an OLR model will be conducted with investment levels as dependent variable and the OXT condition as independent variable (value equals 1 when on OXT and 0 when placebo). Levels of nasal obstruction (NOSE80) will be included as a covariate that could potentially limit the impact of intranasal OXT administration4. The analysis will be performed using the *MASS* package65 implemented in *R*66. If the proportional odds assumption of the OLR model is violated, we will conduct and report the results of three separate BL models (I: 0=0 and 4/8/12=1; II: 0/4=0 and 8/12=1; III: 0/4/8=0 and 12=1) with no correction for post hoc comparisons. If there is no (statistically significant) effect of OXT, the odds ratios will not differ significantly from 1.

***Hypothesis 1b***

Hypothesis 1a will be supplemented with equivalence testing using the *TOSTER* package91 implemented in *R*66. Here, the aim is to assess, in the event that the OXT group, relative to placebo group, shows a significant increase in monetary investments, whether OXT’s effect on interpersonal trust is large enough to be considered a meaningful psychological finding40,47. We will set the upper bound (ΔU)t to *d*=.33, and the lower bound (ΔL) to *d*=-.33 (testing one-sided). Although at first glance this may seem like a rather wide range, it is not our aim to demonstrate equivalence for the smallest possible effect sizes. Rather, we aim to demonstrate equivalence for a range of effect sizes that are smaller than effect sizes commonly encountered in psychological research92. Moreover, this effect size is in the range of the mean effect size reported by previous studies that have examined OXT’s effects in trust-related contexts93-95.

In case the 90% confidence interval of the effect of OXT on investments falls within the range of the upper and lower bound of the pre-defined interval, we have evidence to conclude that the OXT and placebo group are equivalent (where equivalence would imply that the OXT effect is not meaningful enough to further pursue in lab-based studies).

***Hypothesis 2***

To investigate if the effect of OXT on investments (dependent variable) decreases with increasing trust propensity, Hypothesis 1a will be re-run with IGTS scores36 as an additional interaction factor. In contrast to a median split on IGTS that was used in a previous replication4, this variable will be treated as continuous in the present study.

***Hypothesis 3***

To investigate if the effect of OXT on investments (dependent variable) decreases with increasing reward sensitivity or decreasing punishment sensitivity, the same approach as for Hypothesis 2 will be used, where IGTS values will be replaced with even and odd sum scores of the SPSRQ-RC values78.

***Pooled analysis***

To investigate if investors on OXT, compared to placebo, do not significantly invest more money in trustees (dependent variable) in a pooled sample of our proposed sample and Declerck et al.’s minimal social condition sample61, the same analysis as in Hypothesis 1 will be conducted. Equivalence testing (*Hypothesis 1b*) will also be repeated in this larger sample.

**Funding**

This study is supported by a Dutch Research Council (NWO) Replication grant (grant no. 401.19.006 to DH). AR, DH, and CK are additionally supported by a Maastricht University Centre for Integrative Neuroscience (CIN) Interfaculty grant.

**Conflict of Interest Statement**

DH has received financial compensation as a consultant for P1vital Products Ltd. These activities were unrelated to the work presented in this manuscript. The other authors declare no competing interests.

**Code availability**

Power simulation codes are available on OSF (https://osf.io/e3sbf?view\_only=a1fc6796bb92424aad28ff10c11fe595).

**Pilot data availability**

Pilot data analyzed in this study are available on OSF (https://osf.io/prx92?view\_only=a1fc6796bb92424aad28ff10c11fe595).

1Jurek, B., & Neumann, I. D. (2018). The oxytocin receptor: from intracellular signaling to behavior. *Physiological Reviews, 98*(3), 1805–908. <https://doi.org/10.1152/physrev.00031.2017>

2Guastella, A. J., & Hickie, I. B. (2016). Oxytocin treatment, circuitry, and autism: a critical review of the literature placing oxytocin into the autism context. *Biological Psychiatry, 79*(3), 234–242. <http://dx.doi.org/10.1016/j.biopsych.2015.06.028>.

3Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature, 435*(7042), 673–676. <https://doi.org/10.1038/nature03701>

4Declerck, C. H., Boone, C., Pauwels, L., Vogt, B., & Fehr, E. (2020). A registered replication study on oxytocin and trust. *Nature Human Behaviour, 4*(6), 646–655. <https://doi.org/10.1038/s41562-020-0878-x>

5Liu, Y., Li, S., Lin, W., Li, W., Yan, X., Wang, X., Pan, X., Rutledge, R. B., & Ma, Y. (2019). Oxytocin modulates social value representations in the amygdala. *Nature Neuroscience, 22*(4), 633–641. <https://doi.org/10.1038/s41593-019-0351-1>

6Russell, J. A., Leng, G., & Douglas, A. J. (2003). The magnocellular oxytocin system, the fount of maternity: adaptations in pregnancy. *Frontiers in Neuroendocrinology, 24*(1), 27–61.<https://doi.org/10.1016/S0091-3022(02)00104-8>

7Arrowsmith, S., & Wray, S. (2014). Oxytocin: its mechanism of action and receptor signalling in the myometrium. *Journal of Neuroendocrinology, 26*(6), 356–369. <https://doi.org/10.1111/jne.12154>

8Uvnäs‐Moberg, K., Widström, A. M., Werner, S., Matthiesen, A. S., & Winberg, J. (1990). Oxytocin and prolactin levels in breast‐feeding women. Correlation with milk yield and duration of breast‐feeding. *Acta Obstetricia et Gynecologica Scandinavica, 69*(4), 301–306. <https://doi.org/10.3109/00016349009036151>

9White‐Traut, R., Watanabe, K., Pournajafi‐Nazarloo, H., Schwertz, D., Bell, A., & Carter, C. S. (2009). Detection of salivary oxytocin levels in lactating women. Developmental Psychobiology: *The Journal of the International Society for Developmental Psychobiology, 51*(4), 367–373. <https://doi.org/10.1002/dev.20376>

10Naber, F., van IJzendoorn, M. H., Deschamps, P., van Engeland, H., & Bakermans-Kranenburg, M. J. (2010). Intranasal oxytocin increases fathers’ observed responsiveness during play with their children: a double-blind within-subject experiment. *Psychoneuroendocrinology, 35*(10), 1583–6. <https://doi.org/10.1016/j.psyneuen.2010.04.007>

11Kohlhoff, J., Eapen, V., Dadds, M., Khan, F., Silove, D., & Barnett, B. (2017). Oxytocin in the postnatal period: Associations with attachment and maternal caregiving. *Comprehensive Psychiatry, 76*, 56–68. <https://doi.org/10.1016/j.comppsych.2017.03.010>

12Schneiderman, I., Zagoory-Shanon, O., Leckman, J. F., & Feldman, R. (2012). Oxytocin during the initial stages of romantic attachment: relations to couples’ interactive reciprocity. *Psychoneuroendocrinology, 37*(8), 1277–85. https://doi.org/10.1016/j.psyneuen.2011.12.021

13Marsh, N., Marsh, A. A., Lee, M. R., & Hurlemann, R. (2021). Oxytocin and the neurobiology of prosocial behavior. *The Neuroscientist, 27*(6), 604–619. [https://doi.org/10.1177/1073858420960111](https://doi.org/10.1177%2F1073858420960111)

14Quintana, D. S., & Guastella, A. J. (2020). An allostatic theory of oxytocin. *Trends in Cognitive Sciences, 24*(7), 515–528. <https://doi.org/10.1016/j.tics.2020.03.008>

15Froemke, R. C., & Young, L. J. (2021). Oxytocin, neural plasticity, and social behavior. *Annual Review of Neuroscience, 44*, 359–381. <https://doi.org/10.1146/annurev-neuro-102320-102847>

16Hurlemann, R., Patin, A., Onur, O. A., Cohen, M. X., Baumgartner, T., Metzler, S., Dziobek, I., Gallinat, J., Wagner, M., Maier, W., & Kendrick, K. M. (2010). Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. *Journal of Neuroscience, 30*(14), 4999–5007. <https://doi.org/10.1523/JNEUROSCI.5538-09.2010>

17Wu, N., Li, Z., & Su, Y. (2012). The association between oxytocin receptor gene polymorphism (OXTR) and trait empathy*. Journal of Affective Disorders, 138*(3), 468–472. <https://doi.org/10.1016/j.jad.2012.01.009>

18Aydogan, G., Furtner, N. C., Kern, B., Jobst, A., Müller, N., & Kocher, M. G. (2017). Oxytocin promotes altruistic punishment. *Social Cognitive and Affective Neuroscience, 12*(11), 1740–7. <https://doi.org/10.1093/scan/nsx101>

19Marsh, N., Scheele, D., Gerhardt, H., Strang, S., Enax, L., Weber, B., Maier, W. & Hurlemann, R. (2015). The neuropeptide oxytocin induces a social altruism bias. *Journal of Neuroscience, 35*(47), 15696–701. <https://doi.org/10.1523/JNEUROSCI.3199-15.2015>

20Ellenbogen, M. A. (2017). Oxytocin and facial emotion recognition. In: Hurlemann, R., & Grinevich, V. (Eds.), *Behavioral pharmacology of neuropeptides: Oxytocin. Current Topics in Behavioral Neurosciences* (pp. 349–374). Springer. <https://doi.org/10.1007/7854_2017_20>

21Lischke, A., Berger, C., Prehn, K., Heinrichs, M., Herpertz, S. C., & Domes, G. (2012). Intranasal oxytocin enhances emotion recognition from dynamic facial expressions and leaves eye-gaze unaffected. *Psychoneuroendocrinology, 37*(4), 475–481. [https://doi.org/10.1016/j.psyneuen.2011.07.015](https://doi.org/10.1016/j.psyneuen.2011.07.015" \t "Persistent link using digital object identifier)

22Mierop, A., Mikolajczak, M., Stahl, C., Béna, J., Luminet, O., Lane, A., & Corneille, O. (2020). How can intranasal oxytocin research be trusted? A systematic review of the interactive effects of intranasal oxytocin on psychosocial outcomes. *Perspectives on Psychological Science, 15*(5), 1228–42. [https://doi.org/10.1177/1745691620921525](https://doi.org/10.1177%2F1745691620921525)

23Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U., & Fehr, E. (2008). Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron, 58*(4), 639–650. <https://doi.org/10.1016/j.neuron.2008.04.009>

24Teed, A. R., Han, K., Rakic, J., Mark, D. B., & Krawczyk, D. C. (2019). The influence of oxytocin and vasopressin on men’s judgments of social dominance and trustworthiness: An fMRI study of neutral faces. *Psychoneuroendocrinology,* 106, 252–258. <https://doi.org/10.1016/j.psyneuen.2019.04.014>

25Scheele, D., Wille, A., Kendrick, K. M., Stoffel-Wagner, B., Becker, B., Güntürkün, O., Maier, W., & Hurlemann, R. (2013). Oxytocin enhances brain reward system responses in men viewing the face of their female partner. *Proceedings of the National Academy of Sciences, 110*(50), 20308–13. <https://doi.org/10.1073/pnas.1314190110>

26Mickey, B. J., Heffernan, J., Heisel, C., Peciña, M., Hsu, D. T., Zubieta, J. K., & Love, T. M. (2016). Oxytocin modulates hemodynamic responses to monetary incentives in humans. *Psychopharmacology, 233*(23), 3905–19. <https://doi.org/10.1007/s00213-016-4423-6>

27Ide, J. S., Nedic, S., Wong, K. F., Strey, S. L., Lawson, E. A., Dickerson, B. C., Wald, L. L., La Camera, G., & Mujica-Parodi, L. R. (2018). Oxytocin attenuates trust as a subset of more general reinforcement learning, with altered reward circuit functional connectivity in males. *Neuroimage,*174, 35–43. <https://doi.org/10.1016/j.neuroimage.2018.02.035>

28Yao, S., Zhao, W., Geng, Y., Chen, Y., Zhao, Z., Ma, X., Xu, L., Becker, B., & Kendrick, K. M. (2018). Oxytocin facilitates approach behavior to positive social stimuli via decreasing anterior insula activity. *International Journal of Neuropsychopharmacology, 21*(10), 918–925. <https://doi.org/10.1093/ijnp/pyy068>

29Yatawara, C. J., Einfeld, S. L., Hickie, I. B., Davenport, T. A., & Guastella, A. J. (2016). The effect of oxytocin nasal spray on social interaction deficits observed in young children with autism: a randomized clinical crossover trial. *Molecular Psychiatry, 21*(9), 1225–31. <https://doi.org/10.1038/mp.2015.162>

30Shilling, P. D., & Feifel, D. (2016). Potential of oxytocin in the treatment of schizophrenia*. CNS drugs, 30*(3), 193–208. <https://doi.org/10.1007/s40263-016-0315-x>

31Gossen, A., Hahn, A., Westphal, L., Prinz, S., Schultz, R. T., Gründer, G., & Spreckelmeyer, K. N. (2012). Oxytocin plasma concentrations after single intranasal oxytocin administration–a study in healthy men. *Neuropeptides, 46*(5), 211–215. <https://doi.org/10.1016/j.npep.2012.07.001>

32Young, L. J., & Flanagan-Cato, L. M. (2012). Editorial comment: oxytocin, vasopressin and social behavior. *Hormones and Behavior, 61*(3), 227–229. <https://doi.org/10.1016/j.yhbeh.2012.02.019>

33Mikolajczak, M., Pinon, N., Lane, A., de Timary, P., & Luminet, O. (2010). Oxytocin not only increases trust when money is at stake, but also when confidential information is in the balance. *Biological Psychology, 85*(1), 182–184. <https://doi.org/10.1016/j.biopsycho.2010.05.010>

34Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review, 23*(3), 393–404. <https://doi.org/10.5465/amr.1998.926617>

35Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior, 10*(1), 122–142. <https://doi.org/10.1006/game.1995.1027>

36Yamagishi, T., Akutsu, S., Cho, K., Inoue, Y., Li, Y., & Matsumoto, Y. (2015). Two-component model of general trust: Predicting behavioral trust from attitudinal trust. *Social Cognition, 33*(5), 436–458. <https://doi.org/10.1521/soco.2015.33.5.436>

37Lane, A., Mikolajczak, M., Treinen, E., Samson, D., Corneille, O., de Timary, P., & Luminet, O. (2015). Failed replication of oxytocin effects on trust: the envelope task case. *PloS One, 10*(9), e0137000. <https://doi.org/10.1371/journal.pone.0137000>

38Ebert, A., Kolb, M., Heller, J., Edel, M. A., Roser, P., & Brüne, M. (2013). Modulation of interpersonal trust in borderline personality disorder by intranasal oxytocin and childhood trauma. *Social Neuroscience, 8*(4), 305–313. <https://doi.org/10.1080/17470919.2013.807301>

39Nave, G., Camerer, C., & McCullough, M. (2015). Does oxytocin increase trust in humans? A critical review of research. *Perspectives on Psychological Science, 10*(6), 772–789. [https://doi.org/10.1177/1745691615600138](https://doi.org/10.1177%2F1745691615600138)

40Tabak, B. A., Teed, A. R., Castle, E., Dutcher, J. M., Meyer, M. L., Bryan, R., Irwin, M. R., Liebermann, M. D. & Eisenberger, N. I. (2019). Null results of oxytocin and vasopressin administration across a range of social cognitive and behavioral paradigms: Evidence from a randomized controlled trial. *Psychoneuroendocrinology,*107, 124–132. [https://doi.org/10.1016/j.psyneuen.2019.04.019](https://doi.org/10.1016/j.psyneuen.2019.04.019" \t "Persistent link using digital object identifier)

41Lane, A., Luminet, O., Nave, G., & Mikolajczak, M. (2016). Is there a Publication Bias in Behavioural Intranasal Oxytocin Research on Humans? Opening the File Drawer of One Laboratory. *Journal of Neuroendocrinology, 28*(4). <https://doi.org/10.1111/jne.12384>

42Quintana, D. S. (2020). Most oxytocin administration studies are statistically underpowered to reliably detect (or reject) a wide range of effect sizes. *Comprehensive Psychoneuroendocrinology, 4*, 100014. <https://doi.org/10.1016/j.cpnec.2020.100014>

43Pereira, T. V., & Ioannidis, J. P. A. (2011). Statistically significant meta-analyses of clinical trials have modest credibility and inflated effects. *Journal of Clinical Epidemiology, 64*(10), 1060–9. <https://doi.org/10.1016/j.jclinepi.2010.12.012>

44Walum, H., Waldman, I. D., & Young, L. J. (2016). Statistical and methodological considerations for the interpretation of intranasal oxytocin studies. *Biological Psychiatry, 79*(3), 251–257. <https://doi.org/10.1016/j.biopsych.2015.06.016>

45Button, K. S., Ioannidis, J., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14*(5), 365–376. <https://doi.org/10.1038/nrn3475>

46Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science, 1*(2), 259–269. [https://doi.org/10.1177/2515245918770963](https://doi.org/10.1177%2F2515245918770963)

47Quintana, D. S. (2018). Revisiting non-significant effects of intranasal oxytocin using equivalence testing. *Psychoneuroendocrinology, 87*, 127–130. <https://doi.org/10.1016/j.psyneuen.2017.10.010>

48Kim, Y. R., Oh, S. M., Corfield, F., Jeong, D. W., Jang, E. Y., & Treasure, J. (2014). Intranasal oxytocin lessens the attentional bias to adult negative faces: a double blind within-subject experiment. *Psychiatry Investigation, 11*(2), 160–166. <https://doi.org/10.4306/pi.2014.11.2.160>

49Zhuang, Q., Zhu, S., Yang, X., Zhou, X., Xu, X., Chen, Z., Lan, C., Zhao, W., Becker, B., Yao, S., & Kendrick, K. M. (2021). Oxytocin-induced facilitation of learning in a probabilistic task is associated with reduced feedback-and error-related negativity potentials. *Journal of* *Psychopharmacology, 35*(1), 40–49. <https://doi.org/10.1177/0269881120972347>

50Xu, X., Li, J., Chen, Z., Kendrick, K. M., & Becker, B. (2019). Oxytocin reduces top-down control of attention by increasing bottom-up attention allocation to social but not non-social stimuli–a randomized controlled trial. *Psychoneuroendocrinology, 108*, 62–69. [https://doi.org/10.1016/j.psyneuen.2019.06.004](https://doi.org/10.1016/j.psyneuen.2019.06.004" \t "Persistent link using digital object identifier)

51Davis, C., Zai, C. C., Adams, N., Bonder, R., & Kennedy, J. L. (2019). Oxytocin and its association with reward-based personality traits: A multilocus genetic profile (MLGP) approach. *Personality and Individual Differences, 138*, 231–236. <https://doi.org/10.1016/j.paid.2018.09.002>

52Roberts, B. Z., Young, J. W., He, Y. V., Cope, Z. A., Shilling, P. D., & Feifel, D. (2019). Oxytocin improves probabilistic reversal learning but not effortful motivation in Brown Norway rats. *Neuropharmacology,* 150, 15–26. <https://doi.org/10.1016/j.neuropharm.2019.02.028>

53Cavalli, J., Ruttorf, M., Pahi, M. R., Zidda, F., Flor, H., & Nees, F. (2017). Oxytocin differentially modulates pavlovian cue and context fear acquisition. *Social Cognitive and Affective Neuroscience, 12*(6), 976–983. <https://doi.org/10.1093/scan/nsx028>

54Burgstaller, J., Paulus, M., & Pfundmair, M. (2019). Oxytocin promotes action prediction. *Hormones and Behavior,* 107, 46–48. <https://doi.org/10.1016/j.yhbeh.2018.09.004>

55Mislin, A., Williams, L. V., & Shaughnessy, B. A. (2015). Motivating trust: Can mood and incentives increase interpersonal trust? *Journal of Behavioral and Experimental Economics, 58*, 11–19. <https://doi.org/10.1016/j.socec.2015.06.001>

56Delgado, M. R., Frank, R. H., & Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience, 8*(11), 1611–8. https://doi.org/10.1038/nn1575

57Fehr, E., & Camerer, C. F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends in Cognitive Sciences, 11*(10), 419–427. <https://doi.org/10.1016/j.tics.2007.09.002>

58Bryan, C. J., Tipton, E., & Yeager, D. S. (2021). Behavioural science is unlikely to change the world without a heterogeneity revolution. *Nature Human Behaviour*, *5*(8), 980–989. <https://doi.org/10.1038/s41562-021-01143-3>

59Bartz, J. A., Zaki, J., Bolger, N., & Ochsner, K. N. (2011). Social effects of oxytocin in humans: context and person matter. *Trends in Cognitive Sciences, 15*(7), 301–309. [https://doi.org/10.1016/j.tics.2011.05.002](https://doi.org/10.1016/j.tics.2011.05.002" \t "Persistent link using digital object identifier)

60Nawijn, L., van Zuiden, M., Koch, S. B., Frijling, J. L., Veltman, D. J., & Olff, M. (2016). Intranasal oxytocin enhances neural processing of monetary reward and loss in post-traumatic stress disorder and traumatized controls. *Psychoneuroendocrinology,* 66, 228–237. <https://doi.org/10.1016/j.psyneuen.2016.01.020>

61Declerck, C. (2019, September 23). Oxytocin and Trust. <https://osf.io/jkcv5/?view_only=c647a145f38d4717ac4d750d04e1e222>

62Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree – An open-source platform for laboratory, online, and field experiments*. Journal of Behavioral and Experimental Finance*, 9, 88–97. <https://doi.org/10.1016/j.jbef.2015.12.001>

63Calin-Jageman, R. J., & Cumming, G. (2019). The new statistics for better science: Ask how much, how uncertain, and what else is known. *The American Statistician, 73*:sup1, 271–280. <https://doi.org/10.1080/00031305.2018.1518266>

64Calin-Jagerman, R. (2019, January 10). Kosfeld et al. 2005 -- data and re-analysis. <https://osf.io/54n9q/>

65Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S (4th ed.).* Springer.

66R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <https://www.R-project.org/>.

67Brant, R. (1990). Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression. *Biometrics, 46*(4), 1171–8. <https://doi.org/10.2307/2532457>

68Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta‐analysis. *Statistics in Medicine, 19*(22), 3127–31. [https://doi.org/10.1002/1097-0258(20001130)19:22<3127::AID-SIM784>3.0.CO;2-M](https://doi.org/10.1002/1097-0258(20001130)19:22%3c3127::AID-SIM784%3e3.0.CO;2-M)

69Quintana, D. S. (2022, July 11). A guide for calculating study-level statistical power for meta-analyses. <https://doi.org/10.31219/osf.io/js79t>

70Liccardi, G., Bilò, M. B., Mauro, C., Salzillo, A., Piccolo, A., D'Amato, M., & D'Amato, G. (2013). Oxytocin: a likely underestimated risk for anaphylactic reactions in delivering women sensitized to latex. *Annals of Allergy, Asthma & Immunology, 110*(6), 465–466. <https://doi.org/10.1016/j.anai.2013.03.014>

71Winterton, A., Westlye, L. T., Steen, N. E., Andreassen, O. A., & Quintana, D. S. (2021). Improving the precision of intranasal oxytocin research. *Nature Human Behaviour, 5*(1), 9–18. https://doi.org/10.1038/s41562-020-00996-4

72Lee, M. R., Shnitko, T. A., Blue, S. W., Kaucher, A. V., Winchell, A. J., Erikson, D. W., Grant, K. A., & Leggio, L. (2020). Labeled oxytocin administered via the intranasal route reaches the brain in rhesus macaques. *Nature Communications, 11*(1), 1–10. https://doi.org/10.1038/s41467-020-15942-1

73van der Aart, J., Golla, S. S., van der Pluijm, M., Schwarte, L. A., Schuit, R. C., Klein, P. J., Metaxas, A., Windhorst, A. D., Boellaard, R., Lammertsma, A. A., & van Berckel, B. N. (2018). First in human evaluation of [18 F] PK-209, a PET ligand for the ion channel binding site of NMDA receptors. *EJNMMI research, 8*(1), 1–12. https://doi.org/10.1186/s13550-018-0424-2

74Labuschagne, I., Phan, K. L., Wood, A., Angstadt, M., Chua, P., Heinrichs, M., Stout, J. 71 C., & Nathan, P. J. (2010). Oxytocin attenuates amygdala reactivity to fear in generalized social anxiety disorder. *Neuropsychopharmacology, 35*(12), 2403–13. <https://doi.org/10.1038/npp.2010.123>

75Luo, R., Xu, L., Zhao, W., Ma, X., Xu, X., Kou, J., Gao, Z., Becker, B., & Kendrick, K. M. (2017). Oxytocin facilitation of acceptance of social advice is dependent upon the perceived trustworthiness of individual advisors. *Psychoneuroendocrinology,* 83, 1–8. <https://doi.org/10.1016/j.psyneuen.2017.05.020>

76Quintana, D. S., Westlye, L. T., Hope, S., Nærland, T., Elvsåshagen T., Dørum, E., Rustan, Ø, Valstad, M., Rezvaya, L., Lishaugen, H., Stensønes, E., Yaqub, S., Smerud, K. T., Mahmoud, R. A., Djupesland, P. G., & Andreassen, O. A. (2017). Dose-dependent social-cognitive effects of intranasal oxytocin delivered with novel Breath Powered device in adults with autism spectrum disorder: a randomized placebo-controlled double-blind crossover trial. *Translational Psychiatry, 7*(5), e1136. <https://doi.org/10.1038/tp.2017.103>

77Guastella, A. J., Hickie, I. B., McGuinness, M. M., Otis, M., Woods, E. A., Disinger, H. M., Chan, H.-K., Chen, T. F., & Banati, R. B. (2013). Recommendations for the standardisation of oxytocin nasal administration and guidelines for its reporting in human research. *Psychoneuroendocrinology, 38*(5), 612–625. <https://doi.org/10.1016/j.psyneuen.2012.11.019>

78Conner, B. T., Rahm-Knigge, R. L., & Jenkins, A. L. (2018). Revision and clarification of the sensitivity to punishment sensitivity to reward questionnaire. *Personality and Individual Differences*, 121, 31–40. <https://doi.org/10.1016/j.paid.2017.09.016>

79Fisman, R., Jakiela, P., & Kariv, S. (2017). Distributional preferences and political behavior. *Journal of Public Economics, 155,* 1–10. <https://doi.org/10.1016/j.jpubeco.2017.08.010>

80Lipan, M. J., & Most, S. P. (2013). Development of a severity classification system for subjective nasal obstruction. *JAMA Facial Plastic Surgery, 15*(5), 358–361. [http://doi.org/10.1001/jamafacial.2013.344](https://doi.org/10.1001/jamafacial.2013.344)

81Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief “red flags” for autism screening: the short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child & Adolescent Psychiatry, 51*(2), 202–212. <https://doi.org/10.1016/j.jaac.2011.11.003>

82Van Lange, P. A. (2000). Beyond self-interest: A set of propositions relevant to interpersonal orientations. *European Review of Social Psychology, 11*(1), 297–331. <https://doi.org/10.1080/14792772043000068>

83Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology, 67*(6), 1063–78. [https://doi.org/10.1037/0022-3514.67.6.1063](https://doi.apa.org/doi/10.1037/0022-3514.67.6.1063)

84Coutlee, C. G., Politzer, C. S., Hoyle, R. H., & Huettel, S. A. (2014). An Abbreviated Impulsiveness Scale (ABIS) Constructed through Confirmatory Factor Analysis of the BIS-11. *Archives of Scientific Psychology, 2*(1), 1–12. https://doi.org/10.1037/arc0000005

85Steyer, R., Schwenkmezger, P., Notz, P., & Eid, M. (1997). *Der Mehrdimensionale Befindlichkeitsfragebogen MDBF [Multidimensional mood questionnaire]*. Hogrefe.

86Lee, K., & Ashton, M. C. (2018). Psychometric properties of the HEXACO-100. *Assessment, 25*(5), 543–556. [https://doi.org/10.1177/1073191116659134](https://doi.org/10.1177%2F1073191116659134)

87Falk, A., Becker, A., Dohmen, T. J., Huffman, D., & Sunde, U. (2016). The preference survey module: A validated instrument for measuring risk, time, and social preferences. *Netspar Discussion Paper No. 01/2016-003*. <https://dx.doi.org/10.2139/ssrn.2725874>

88Dieleman, G. C., van der Ende, J., Verhulst, F. C., & Huizink, A. C. (2010). Perceived and physiological arousal during a stress task: can they differentiate between anxiety and depression?. *Psychoneuroendocrinology, 35*(8), 1223–34. [https://doi.org/10.1016/j.psyneuen.2010.02.012](https://doi.org/10.1016/j.psyneuen.2010.02.012" \t "Persistent link using digital object identifier)

89Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology, 63*(4), 596–612. [https://doi.org/10.1037/0022-3514.63.4.596](https://psycnet.apa.org/doi/10.1037/0022-3514.63.4.596)

90Teixeira, R. U. F., Zappelini, C. E. M., Oliveira, L. G., Basile, L. C. G., & da Costa, E. A. (2011). Correlation Between the Peak Nasal Inspiratory Flow and the Visual Analogue Scale Before and After Using a Nasal Decongestant. *International Archives of Otorhinolaryngol, 15(*2), 156–162. http://www.arquivosdeorl.org.br/additional/acervo\_eng.asp?id=758

91Lakens, D. (2017). Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science, 8*(4), 355–362. <https://doi.org/10.1177/1948550617697177>

92Schäfer, T., & Schwarz, M. A. (2019). The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. *Frontiers in Psychology*, *10*(813). https://doi.org/10.3389/fpsyg.2019.00813

93Mikolajczak M., Gross J. J., Lane A., Corneille O., de Timary P., Luminet O. (2010). Oxytocin makes people trusting, not gullible. *Psychological Science, 21*, 1072–4.

94Klackl J., Pfundmair M., Agroskin D., Jonas E. (2013). Who is to blame? Oxytocin promotes nonpersonalistic attributions in response to a trust betrayal. *Biological Psychology, 92*, 387–394

95Yao S., Zhao W., Cheng R., Geng Y., Luo L., Kendrick K. M. (2014). Oxytocin makes females, but not males, less forgiving following betrayal of trust. *The International Journal of Neuropsychopharmacology, 17*, 1785–92

**Appendix A**

General and specific investor trust game instructions (based on Declerck et al.4 and adjusted to the purpose of this study)

**General instructions**

You are now participating in a scientific study.

If you read the following instructions carefully, you can earn money, depending on your decisions and those of other participants. For this reason, it is very important that you read these instructions carefully.

**Please note that you are not permitted to communicate with other participants in any way during this study.** Should you have any questions, please raise your hand to and one of us will come to you to answer your questions in private.

During the study, your income will be calculated in Experimental Money Units (EMU). At the end of the experiment, your EMU earnings will be exchanged into Euro at a rate of:

**1 EMU = 0.75 Euro**.

You will receive your total income via bank transfer after the end of this study The amount of payment will be secret. Thus, no other participant will learn how much you earned.

All participants will make their decisions via the computer. During the study, you will be **randomly** **matched with other participants** **who you just met**. Neither before, nor after the study will you learn the exact identity of the participants you are matched with. In the same way, the matched participants will not be informed about your identity.

We describe the precise course of the study on the next pages.

**The Study**

There are two different participant roles: Participant A and Participant B who will interact with each other.

**You are in the role of Participant A**

Both participants, A and B, receive an initial endowment of 12 EMU at the beginning of the interaction. Then, you as Participant A make a single decision to transfer 0, 4, 8, or 12 EMU of your initial endowment to Participant B. Each EMU you transfer to Participant B will be tripled. That is, if you

* transfer 0 EMU, then Participant B receives 0 EMU in addition to his initial endowment;
* transfer 4 EMU, then Participant B receives 12 EMU, in addition to his initial endowment;
* transfer 8 EMU, then Participant B receives 24 EMU, in addition to his initial endowment;
* transfer 12 EMU, then Participant B receives 36 EMU, in addition to his initial endowment

Participant B will be informed of your decision after you have made it. Participant B then has to decide how much of his endowment (initial endowment of 12 EMU + additional EMU received from you) he will transfer back to you. EMU transferred back will not be tripled, meaning that you will receive exactly the number of EMU that Participant B transfers back to you.

Consider this arbitrary example: in case you transfer 4 EMU to Participant B, then you are left with 8 EMU (12 – 4 = 8) and he has 24 EMU (4 × 3 = 12, plus 12 initial endowment) available. He can then transfer back to you any whole EMU amount between 0 and 24. If Participant B, for example, transfers 7 EMU back to you, then **you as** **Participant A** **earn** 15 EMU (7 EMU transferred back by participant B to you, plus 8 remaining EMU of your initial endowment). In this example, **Participant B earns** 17 EMU (24 – 7).

Thus, depending on your transfer and Participant B's decision, both participants' income is determined as follows:

**You as Participant A earn:**

12 – your transfer to Participant B + Participant B’s back transfer to you.

**Participant B earns:**

12 + (3 × your transfer to Participant B) – Participant B’s back transfer to you.

To let you practice calculating your income, here are two more examples for you to consider.

**Example 1:** Suppose you as Participant A transfer 8 EMU to Participant B and this participant transfers 12 EMU back to you. How much do you and Participant B earn in this case?

**You as** **Participant A** **earn**  \_\_\_\_ EMU

**Participant B earns** \_\_\_\_ EMU

**Example 2**: Suppose that you, as in the first example, transfer 8 EMU to Participant B. Participant B now decides to transfer 2 EMU back to you. How much do you and Participant B earn in this case?

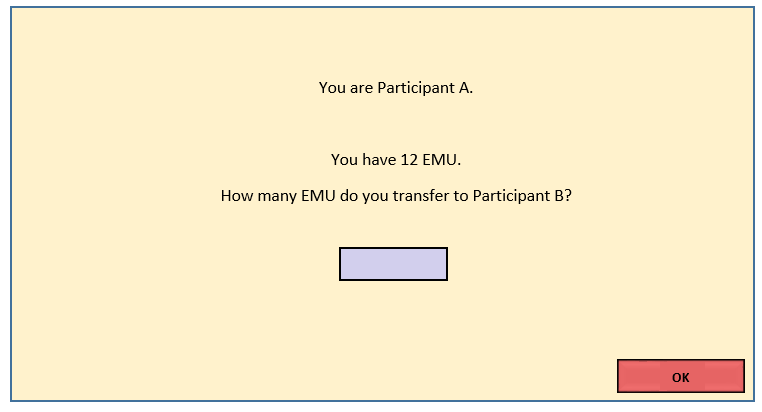
**You as Participant A** **earn** \_\_\_\_ EMU

**Participant B** **earns** \_\_\_\_ EMU

Note: These are two arbitrary examples. Naturally, many other combinations of decisions are possible, leaving you with either more, or less, than your initial endowment.

*Please raise your hand when you have answered both questions. The room supervisor will check your answers. Only after your answers have been checked and are correct, can you proceed to the next page.*

When making your decision you will see a screen like the one shown below. You will enter your transfer decision using the keyboard and confirm your decision with clicking the OK button.



**IMPORTANT**

Remember that as Participant A you can choose between 0, 4, 8, or 12 EMU for your transfer. Participant B can choose any integer amount for the back transfer with the minimum amount being zero and the maximum amount being the number of EMU he has available after your transfer decision.

The Participant B with whom you are matched is **determined randomly from among the participants** **whom you just met**. You will not be informed with which one of these participants you are matched. You only know that it is one of the participants whom you just met. In this sense, the interaction is anonymous, meaning that neither you nor Participant B will be informed about the other participant's identity during or after the experiment.

You will learn of the **decision** of the Participant B with whom you are matched **at the end of this session**. You will then also learn your earnings. How much you earn depends on both your decision and the decision of the Participant B with whom you are matched.

**Appendix B**

Specific trustee trust game instruction (based on Declerck et al.4 and adjusted to the purpose of this study)

**You are now in the role of Participant B**

You now have an additional opportunity to earn money. For the following decision you take on the role of Participant B, and you are matched with a participant A. **This participant A is another person than the participant B you have been matched with before.**

In this part of the study, everything is the same as before with one important difference: **you now have the role of Participant B** and you are **randomly matched with another participant in the role of Participant A**.

Both participants, A and B, receive an initial endowment of 12 EMU at the beginning of the interaction. Then Participant A can transfer 0, 4, 8, or 12 EMU of his initial endowment to you as Participant B. Each EMU which Participant A transfers to you will be tripled. If, for example, A transfers 4 euro, you will receive 12 euro, in addition to your initial endowment. That is, if Participant A

* transfers 0 EMU, then you receive 0 EMU in addition to your initial endowment;
* transfers 4 EMU, then you receive 12 EMU, in addition to your initial endowment;
* transfers 8 EMU, then you receive 24 EMU, in addition to your initial endowment;
* transfers 12 EMU, then you receive 36 EMU, in addition to your initial endowment

You will be informed of Participant A's decision after he has made it. You then have the possibility of transferring back any amount of your EMU to Participant A. EMU transferred back will not be tripled, meaning that Participant A will receive exactly the number of EMU that you transfer.

Consider this arbitrary example: in case Participant A transfers 4 EMU to you as Participant B, then you have 24 EMU (4 × 3 = 12 plus 12 initial endowment) available and Participant A has 8 EMU left (12 – 4 = 8). You can then transfer back any whole EMU amount between 0 and 24. If you, for example, transfer 7 EMU back, then **you as Participant B** earn 17 EMU (24 – 7). **Participant A** then **earns** 15 EMU (7 EMU transferred back by you + 8 remaining EMU of his initial endowment).

To provide another example: in case Participant A transferred 8 EMU, you have 36 EMU (8 × 3 = 24 plus 12 initial endowment) available; you can then transfer back between 0 and 36 EMU. If you, for example, transfer back 12 EMU, then **you as Participant B** earn 24 EMU (36 – 12). **Participant A** then **earns** 16 EMU (12 + 4 remaining euro of his initial endowment).

Note: These are two arbitrary examples. Naturally, many other combinations of decisions are possible, leaving you and Participant A with more or less earnings.

Depending on Participant A's decision and the amount of your back transfer, both participants' income is determined as follows:

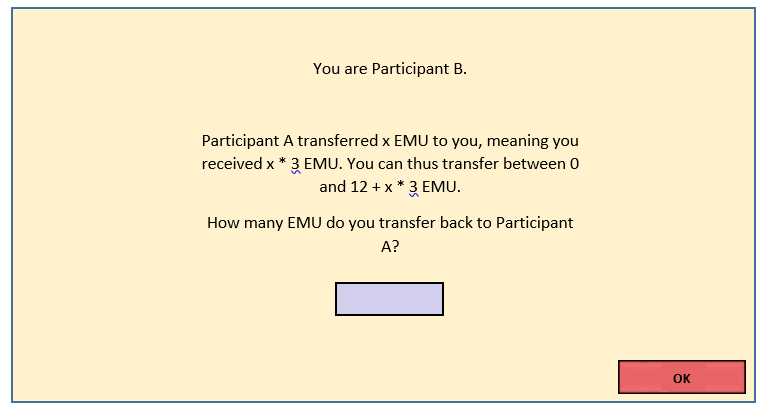
**Participant A earns:**

12 – Participant A's transfer to you + your back transfer to Participant A.

**You as Participant B earn:**

12 + (3 × Participant A's transfer to you) – your back transfer to Participant A.

When making your decision you will see a screen like the one shown below. You will enter your back transfer decision using the keyboard and confirm your decision with clicking the OK button.



**IMPORTANT**

Participant A can transfer 0, 4, 8, or 12 EMU to you. You, as participant B, can choose any integer amount for the back transfer with the minimum amount being zero and the maximum amount being the number of EMU you have available after Participant A’s transfer decision.

The Participant A with whom you are matched is **determined randomly from among the participants whom you just met**. Importantly, you will be matched with a different participant than in the previous situation, and you will not be informed with which one you will be matched. You only know that it is one of the participants whom you met. In this sense, the interaction is anonymous, meaning that neither you nor Participant B will be informed about the other participant's identity during or after the experiment.

**Appendix C**

Guidelines for oxytocin administration, based on recommendations by Guastella et al., 201377

|  |  |
| --- | --- |
| **1** | If necessary, clear your nose from any obstruction (box of tissues provided). |
| **2** | Prime the bottle and complete a test spray in the air. |
| **3** | Sit comfortably and keep the head in an upright position. |
| **4** | Close one nostril with one finger while administering the spray to the other nostril. |
| **5** | Insert bottle 1 cm into the nostril and keep the tip of the bottle at a 45 degree angle into the nose. Aim towards the upper lateral part of the nose (and not towards the middle of the nose). |
| **6** | Upon delivery, inhale and breathe in lightly. Do no sniff exaggeratedly. |
| **7** | Alternate administration between nostrils. Allow time between each re-administration to the same nostril at least 15 seconds. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Question | Hypothesis | Sampling plan | Analysis Plan | Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis | Interpretation given different outcomes | Theory that could be shown wrong by the outcomes |
| Can we replicate the finding from Kosfeld et al. (2005)3: Does intranasal administration of oxytocin, compared to a placebo, increase trusting behavior? | H1a (main hypothesis): Investors on oxytocin, compared to placebo, do not significantly invest more money in trustees. | We base our sample size on an a priori power analysis. For this we used *Figure 2a* from Kosfeld et al.3 and the corresponding reconstructed dataset63 for missing placebo values.  To derive probabilities for the investment amounts of 0, 4, 8, and 12 EMU, we collapsed probabilities in line with Kosfeld et al.3. We then extracted probabilities from both the oxytocin and placebo condition and simulated 1000 random data sets to derive the empirical power for detecting a significant group difference based on an ordinal logistic regression.  Using the probabilities reported in Kosfeld et al.3 and with *α* set to .02 (testing one-sided), a sample of *n*=220 would yield a power of .99 to detect an increase in investments following oxytocin administration. | We will conduct an ordinal logistic regression model with investment levels as dependent variable and the oxytocin condition as independent variable (value equals 1 when on oxytocin, and 0 when on placebo). Levels of nasal obstruction assessed will be included as a covariate that could potentially limit the impact of intranasal oxytocin administration.  If the proportional odds assumption of the ordinal logistic regression model is violated, we will conduct and report the results of three separate binary logistic models (I:0=0 and 4/8/12=1; II: 0/4=0 and 8/12=1; III: 0/4/8=0 and 12=1) with no correction for post hoc comparisons. | We believe that one-sided hypothesis testing (investments oxytocin > investments placebo) at *p*<0.02 is justified since this is already a very strict threshold alpha which we consistently use for testing our replication-related hypotheses 1-3. That is, an alpha of .02 is close to a stringent Bonferroni-corrected alpha of .05 corrected for the number of hypotheses. | If the odds ratios will differ significantly from 1, participants on oxytocin vs. on placebo differ in terms of their trusting behavior (measured by investments in the trust game) and we will have evidence that oxytocin has an effect on trusting behavior. | Theory: oxytocin is a trust-enhancing hormone.  In case we find no evidence that oxytocin increases trust as measured by the trust game (both H1a and pooled analysis), we will have found no supporting evidence for the theory that oxytocin serves a trust-promoting role. |
| In terms of effects of oxytocin administration on trusting behavior, how can we differentiate between true null effects and significant effects that are too small to be considered meaningful? | H1b: Even if, contrary to the prediction, we would observe a significant difference between the OXT and placebo group, we expect it to be significantly smaller than a predefined effect size of interest | The bounds of the equivalence interval will be set to the effect size that the original study3 would have 33% power to detect (*n*=29 for both the oxytocin and placebo group, Mann-Whitney-U-test, one-sided). Consequently, ΔU will be set to d=.33 and ΔL will be set to *d*=-.33.  Equivalence testing will be done according to Lakens85. | For our proposed equivalence tests, we settled on an effect size interval that contains effect sizes that are considerably smaller than the originally reported effect size (i.e., ~2-3 times smaller) and that would allow us to conclude that oxytocin effects on trusting behavior may a) be difficult to detect in lab-based experiments and b) less meaningful than originally thought. | In case the 90% confidence interval of the effect of oxytocin on investments falls within the range of the upper and lower bound of the pre-defined interval, we have evidence to conclude that the oxytocin and placebo group are equivalent ( (where equivalence would imply that the oxytocin effect is not meaningful enough to pursue in future lab-based studies). | In case we find that a difference between the oxytocin and placebo group is significantly smaller than a predefined effect size of interest (i.e., by conducting equivalence testing), this effect will be difficult to pursue/observe in lab-based studies because it is very small. |
| Is there a specific subpopulation, in particular people with low baseline trust propensity, whose investment behavior may be especially sensitive to oxytocin manipulations? | H2: We expect the effect of oxytocin on investments to decrease with increasing trust propensity of the participants. | We will use the same approach as in H1a with mean values from *the Inclusive Generalised Trust Scale* as an additional interaction factor. | We believe that testing for the presence of this group (oxytocin, placebo) by questionnaire score interaction at *p*<0.02 is justified since this is already a very strict threshold alpha which we consistently use for testing our replication-related hypotheses 1-3. That is, an alpha of .02 is close to a stringent Bonferroni-corrected alpha of .05 corrected for the number of hypotheses. . We will report this result in combination with the effect size, so that readers can decide upon the relevance of this result. | If the odds ratios for the interaction term do not differ significantly from 1, we will have no evidence that effects of oxytocin on trusting behavior are influenced by baseline trust propensity. | Theory: oxytocin is a trust-enhancing hormone.  In case we find evidence that the effect of oxytocin on trust measured by the trust game will be dependent on the individual (baseline) trust propensity of participants, this would highlight potential boundary conditions (related to character traits) of the original theory. |
| Is there a specific subpopulation, in particular people with low baseline low reward sensitivity or high punishment sensitivity, whose investment behavior may be especially sensitive to oxytocin manipulations? | H3: We expect the effect of oxytocin on investments to decrease with increasing reward sensitivity and decreasing punishment sensitivity. | We will use the same approach as in H1a with *z*-scored even and odd sum scores of the *Sensitivity to Punishment and Reward Questionnaire – Revised and Clarified* as additional interaction factors. | We believe that testing for the presence of this group (oxytocin, placebo) by questionnaire score interaction at *p*<0.02 uncorrected is justified since this is already a very strict threshold alpha which we consistently use for testing our replication-related hypotheses 1-3. That is, an alpha of .02 is close to a stringent Bonferroni-corrected alpha of .05 corrected for the number of hypotheses. We will report this result in combination with the effect size, so that readers can decide upon the relevance of this result. | If the odds ratios for the interaction term do not differ significantly from 1, we will have no evidence that effects of oxytocin on trusting behavior are influenced by baseline reward and/or punishment sensitivity. | Theory: oxytocin is a trust-enhancing hormone.  In case we find evidence that the effect of oxytocin on trust measured by the trust game will be dependent on the individual (baseline) reward and punishment sensitivity of participants, this would highlight potential boundary conditions (related to character traits) of the original theory. |
| What effect sizes can be expected when using intranasal oxytocin administration to modulate trusting behavior? | Pooled analysis: Pooling the to-be-collected data with the minimal social contact sample data of Declerck et al.61 to increase power to find a small effect of OXT on intrapersonal trust. | Mixed results from previous oxytocin-trust (replication) studies suggest that the true effect regarding an increase in trusting behavior following oxytocin administration is smaller than the effect size of Cohen’s *d* of .51 (corresponding to an odds ratio of approximately 2.52) reported in the original study by Kosfeld et al.3.  To this end, we will combine our data with the minimal social contact condition sample from a previous large-scale replication61.  We used extracted probabilities for the placebo condition from Kosfeld et al.3 and assumed a true minimal effect size of Cohen’s *d* =.2, which would translate to an odds ratio of 1.44. We then simulated 1000 random datasets to derive power.  With *α* set to .05 (testing one-sided), a combined sample of *N*=546 (220 participants collected as part of this pre-registration) would yield a power of .8 to detect an increase in interpersonal trust following OXT administration. Lowering *α* to .02 (again, one-tailed) would still yield a power of about .68. | H1a will be re-run in the pooled sample; corresponding equivalence testing as in H1b will be repeated as well. | Same rationale as H1a, but with the addition that we increase our power to increase sensitivity to even smaller effects of oxytocin on investments (and, for H1b, that we increase power to assess group equivalence). | Same interpretation as for H1a/H1b. | Same interpretation as for H1a/H1b. |