TRUST AND SOCIAL RISK: NEURO-SCIENTIFIC FOUNDATIONS OF ECONOMIC DECISION-MAKING UNDER STRATEGIC UNCERTAINTY

Inaugural-Dissertation

zur Erlangung des Grades Doctor oeconomiae publicae (Dr. oec. publ.)

an der Ludwig-Maximilians-Universität München

2015

vorgelegt von

Gökhan Aydogan

Referent: Prof. Dr. Martin G. Kocher

Korreferent: Prof. Dr. Florian Englmaier

Promotionsabschlussberatung: 4. November 2015

Datum der mündlichen Prüfung: 27. Oktober 2015

Namen der Berichterstatter: Martin G. Kocher, Florian Englmaier, Michaela Pfundmair

TRUST AND SOCIAL RISK: NEURO-SCIENTIFIC FOUNDATIONS OF ECONOMIC DECISION-MAKING UNDER STRATEGIC UNCERTAINTY

Inaugural-Dissertation
zur Erlangung des Grades Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

2015

vorgelegt von

Gökhan Aydogan

Referent: Prof. Dr. Martin G. Kocher
Korreferent: Prof. Dr. Florian Englmaier
Promotionsabschlussberatung: 4. November 2015

Danksagung

Die Entwicklung und Entstehung dieser Arbeit wären nicht ohne die tatkräftige Hilfe und liebevolle Unterstützung von zahlreichen Kollegen, Freunden und meiner Familie möglich gewesen. Wenn ich alle einzeln nennen würde, bräuchte ich mehrere Seiten, weswegen ich mich an dieser Stelle nur auf einen kleinen Teil der Personen beschränken möchte, die mich auf diesem Lebensabschnitt begleitet haben. Falls ich jemanden nicht ausdrücklich genannt haben sollte, ist dies sicherlich dem begrenzten Rahmen geschuldet, weswegen mir das nachgesehen werden möge.

An erster Stelle möchte ich mich bei meinem Doktorvater Martin Kocher bedanken, der mir die Möglichkeit eröffnet hat, in einem immer noch außergewöhnlichen Themenbereich der Ökonomik zu forschen, und mich bei diesem Vorhaben in jeglicher Hinsicht unterstützt hat. Hierbei möchte ich mich auch besonders für die Freiheit bedanken, die er mir im Laufe der Promotion überlassen hat, was unter anderem sicherlich zum Gelingen dieser Dissertation beitrug. Weiterhin möchte ich mich ganz herzlich bei meinem Zweitbetreuer Florian Englmaier bedanken, dessen wertvolle Kommentare und hilfreiche Anregungen ich immer sehr geschätzt habe. Besonderer Dank gebührt auch meiner Drittgutachterin Michaela Pfundmaier für ihre außerordentlich sachkundige, erfahrene und wertvolle Unterstützung.

Des Weiteren bedanke ich mich herzlichst bei meinen Ko-Autoren Sandra Dehning, Nadja Furtner, Bianca Kern, Fabian Loy, Norbert Müller und Peter Zill für die vertrauensvolle und erfolgreiche Zusammenarbeit. An dieser Stelle möchte ich besonderen Dank Andrea Jobst aussprechen, die mit ihrer medizinischen Fachunterstützung und ihrem persönlichen Engagement dieses Projekt erst möglich gemacht hat. Ich bedanke mich auch ganz herzlich für alle Anregungen und Ideen meiner Kollegen am Lehrstuhl sowie an der Fakultät, die ich im Laufe meines Forschungsprojektes erhalten habe.

Zu guter Letzt gebührt mein größter Dank meiner Familie, insbesondere meiner Frau Kerasula und meiner Tochter Melina für ihre liebevolle Geduld und uneingeschränkte Unterstützung, die mir sowohl durch Höhen als auch durch Tiefen meiner Promotion zur Seite standen und mich immer wieder ermutigt haben.

Contents

List of Figures

List of Tables

“What constrains or enables the capacity of human beings to work in groups is not so much the technology, but rather the capacity of the human brain to have and monitor social interactions.”

Nicholas A. Christakis

Preface

Standard economic game theory assumes that agents are fully rational and self-serving, and are able to instantly process an infinite amount of information. However, if all decision-makers are rational and selfish (and this is common knowledge), there is no social risk or strategic uncertainty, as long as there is a unique solution or equilibrium of the interaction structure. Under these assumptions, social risk can exist, for instance, in coordination games in which multiple equilibria lead to the risk of mis-coordination. However, if assumptions regarding rationality or selfish behavior are relaxed, different types of decision makers (e.g., differently bounded rational players or different types of altruistic players) lead to strategic uncertainty in interactions. Such uncertainty is omnipresent on (financial) markets and in any form of bilateral and small-group interaction.

This dissertation intends to shed light on one important aspect of strategic uncertainty. In particular, this work examines the hormonal foundations of human behavior and human decision-making under social risk. To advance and refine economic models, it is indispensable to get an insight of the effects of hormones on behavior in markets and in human interactions. While the impact of some hormones on economic decision making, especially their impact on behavior on financial markets (e.g., trader behavior), have been studied before (Eisenegger, Haushofer, & Fehr, 2011; Sapienza, Zingales, & Maestripieri, 2009), this work explicitly focuses on social risk and on one single hormone for which only limited evidence in an economic context exists: oxytocin.

Oxytocin is a neuropeptide and hormone, which is produced in the hypothalamus and released both into brain and bloodstream (Donaldson & Young, 2008b). Oxytocin, like all neuropeptides, acts as a neuronal messenger in the mammalian brain and influences therefore many neurobehavioral functions (Bethlehem, Baron-Cohen, van Honk, Auyeung, & Bos, 2014).[1] Besides its important physiological functions in parturition and milk letdown, the neuropeptide plays a key role in reproductive behavior including pair bonding, parenting and bonding between mother and infant (MacDonald & MacDonald, 2010).

However, a decade ago, cognitive sciences including psychology and neuro-economics gained a huge interest in oxytocin, as the seminal paper by Kosfeld, Heinrichs, Zak, Fischbacher, and Fehr (2005) suggested that an intranasal application of oxytocin reduces social risk aversion in humans.[2] This led to the view that oxytocin is related to the neurobiological roots of trusting behavior, and established therefore even the reputation of being ‘liquid trust’ (MacDonald & MacDonald, 2010; Zak, Kurzban, & Matzner, 2004).

Following by a huge body of empirical studies, for almost a decade, oxytocin’s reputation gradually shifted to a reputation as being a ‘moral molecule’, since various other pro-social effects have been attributed to oxytocin (Bethlehem et al., 2014; MacDonald & MacDonald, 2010). For instance, it has been shown that an exogenous application of oxytocin improves empathy (Domes, Heinrichs, Michel, Berger, & Herpertz, 2007; Shamay-Tsoory et al., 2013), increases trustworthiness (Zak, Kurzban, & Matzner, 2005) and makes one even more generous (Zak, Stanton, Ahmadi, & Brosnan, 2007).

However, this view was questioned by recent findings that suggested that oxytocin’s effects are not as pro-social as presumed. One of the first contradicting evidence regarding oxytocin’s pro-social reputation was provided by Shamay-Tsoory et al. (2009), as subjects given oxytocin exhibited higher levels of envy and schadenfreude compared to placebo during unequal monetary gain conditions. Additionally, empirical evidence suggests that oxytocin even hinders cooperation and trust when negative social information is present (Declerck, Boone, & Kiyonari, 2010; Declerck, Boone, & Kiyonari, 2013). Recently, it has also been shown that oxytocin biases moral judgments to the favor of one’s in-group with negative consequences for out-group members (De Dreu, Greer, Van Kleef, Shalvi, & Handgraaf, 2011) and additionally, the neuropeptide even motivates parochial altruism and aggressive behavior against competing out-groups (De Dreu et al., 2010). Moreover, oxytocin exhibits even a

---

[1] Within the mammalian brain, oxytocin is mainly synthesized in magnocellular neurons of the hypothalamus that project to the posterior hypophysis. From the hypophysis, oxytocin is released into the bloodstream to act as a hormone by modulating bodily functions. In addition, neurons in the paraventricular nuclei project oxytocin to various brain structures including the hippocampus, amygdala and nucleus accumbens, which contain numerous oxytocin receptors (MacDonald & MacDonald, 2010).

[2] Intranasal administered oxytocin is able to pass the blood-brain barrier and achieve access to the cerebrospinal fluid (CSF) within 30 minutes (Born et al., 2002; Gossen et al., 2012). Accordingly, oxytocin can be easily applied to humans via nasal sprays, which allows studying the role of oxytocin in human behavior with non-invasive measures.

detrimental effect on a widely established norm, i.e. honesty, as it promotes group-serving lying (Shalvi & De Dreu, 2014).

In light of growing evidence that contradicts the notion of being a 'moral molecule', the question arises what might be the common underlying effect that is able to explain likewise all well-documented effects of oxytocin. Besides other hypotheses in the literature, three main approaches have been developed and attracted considerable attention (Bethlehem et al., 2014). For instance, the *social cue hypothesis* suggests that oxytocin might render social information more salient, which could explain the context dependent effects on cooperation and trusting behavior (Bartz, Zaki, Bolger, & Ochsner, 2011). However, this hypothesis is not able to explain the observed effects on ethnocentrism, which persisted even in the absence of intergroup comparisons (De Dreu et al., 2010). Moreover, the *social reward sensitivity hypothesis* suggests that oxytocin's efficacy might be rooted in a stronger reaction to social rewards (Dolen, Darvishzadeh, Huang, & Malenka, 2013; Groppe et al., 2013). However, this hypothesis mainly builds on animal studies and has provided only very scarce evidence regarding humans in incentive-compatible experiments. Additionally, the *social reward hypothesis* is very similar to the *social cues hypothesis* and can therefore be considered merely as an extension of it.

Based on findings that ascribed anxiolytic effects to an exogenous application of oxytocin, a third hypothesis emerged that tries to explain the common underlying effect of oxytocin. According to the *anxiety reduction hypothesis,* oxytocin might reduce anxiety which could constitute a common underlying channel of oxytocin's effect on social behavior (Churchland & Winkielman, 2012). That is, oxytocin-induced lower levels of general anxiety could increase, as a kind of by-product, approaching and trusting behavior. However, there is contradicting evidence to this hypothesis, since subjects given oxytocin exhibited even higher levels of anxiety towards an unpredictable threat (Grillon et al., 2012). Accordingly, additional research on this issue is indispensable, since none of the current hypotheses is able to explain all well-documented effects of oxytocin in a unified approach.

This dissertation constitutes a first step to extend previous findings regarding the effects of oxytocin by employing methods and models from experimental economics in order to infer more general conclusions in this topic. The results of this dissertation may be capable of advancing theories regarding oxytocin's efficacy, and additionally, may inform economics as well. In particular, by shedding light on the neurobiological foundations of economic decision-making, economic models and their predictions could be substantially improved (Camerer, Loewenstein, & Prelec, 2005). Accordingly, this work investigates the neurobiological mechanisms that lead to different types of decision-makers regarding their degree in strategic sophistication (or bounded rationality) and other-regarding preferences. The results of this work may improve our understanding of uncertainty in the form of social risk on markets and in human interactions, in general. The application of oxytocin in interactive decisions is interesting for neurosciences as well, since it indicates new possible clinical treatments in the context of social disorders, like social phobia, schizophrenia or autism (Guastella et al., 2010; Pedersen et al., 2011). Therefore, all studies in this research project were conducted in collaboration with the Clinic for Psychiatry and Psychotherapy at the University of Munich.[3]

This dissertation consists of four essays about the effects of oxytocin on human behavior under strategic uncertainty. In all of them, we employed a fully computerized and anonymous experimental design to ensure that participants could not sense or use any social information of their counterparts other than provided in the experimental tasks. This is of importance, since empirical evidence suggests that participants given oxytocin showed a higher sensitivity to social information in interactive tasks (Bartz et al., 2011; Declerck et al., 2010; Declerck et al., 2013). Moreover, all fully incentivized experiments employed a double-blind placebo-controlled study design to compare economic behavior of participants that received a single dose of intranasal self-administered oxytocin with that of subjects in a control group that received a placebo, containing the same inactive ingredients except for the neuropeptide. This design allowed us to draw causal inferences from an exogenous alteration of brain oxytocin levels to economic behavior (Born et al., 2002; Gossen et al., 2012).

Chapter 1 and 2 explore the role of oxytocin in social norms. While the first chapter examines the influence of oxytocin on norm enforcement, the second chapter analyzes its effects on the adherence to norms when monitoring or sanctioning is not feasible. The third chapter addresses oxytocin's function in social cohesion by analyzing its effects on the endogenous formation of groups. Chapter 4 provides initial evidence on the important role of the oxytocinergic system in strategic reasoning and strategic mentalizing.

In the first chapter, which is joint work with Nadja Furtner, Andrea Jobst, Bianca Kern, Martin Kocher and Norbert Müller, we examine the effect of an exogenous alteration of brain oxytocin levels on the inclination to enforce social norms. In particular, we test whether

[3] In particular, my co-authors from the Clinic for Psychiatry, i.e. Sandra Dehning, Andrea Jobst, Fabian Loy, Norbert Müller and Peter Zill, significantly supported this dissertation by providing profound medical knowledge and by supervising the administration of the nasal sprays in the experimental sessions.

oxytocin is linked to altruistic punishment of defective behavior in a social dilemma situation. This work contributes to the scarce literature that examines the neurological foundations of social norm enforcement (De Quervain et al., 2004; Egas & Riedl, 2008; Fehr & Fischbacher, 2004).

We design an experiment with two different treatments. In the first treatment, participants decide whether to cooperate or defect in a one-shot and simultaneous three person prisoners' dilemma game *without a punishment option and without feedback*. This treatment serves as a baseline, since we are able to elicit general effects of oxytocin on cooperative behavior. In the second treatment, subjects are re-assigned to groups of three in a perfect stranger matching, and played a similar simultaneous three-person prisoners' dilemma game, but this time *with feedback and with a punishment option*. That is, after observing the other two group members' decisions, subjects have the opportunity to reduce the individual payoffs of them by assigning deduction points. This 2 x 2 study design has two main advantages. First, we can examine the effect of oxytocin on the inclination to sanction norm violating behavior. Second, it is possible to compare oxytocin's effects on cooperative behavior in different cooperative institutions, in which a sanctioning mechanisms is either present or absent.

Our results make two main contributions. First, we show that oxytocin substantially increases the inclination and the magnitude of punishment directed towards free riders. The channel through which it operates seems to be an augmented arousal of negative emotions following a norm violation. Our findings extend existing studies that have indicated evidence for oxytocin to stimulate negative emotions (Shamay-Tsoory et al., 2009) or defensive aggression (De Dreu et al., 2010). And second, we find that oxytocin only increases levels of cooperation when a punishment option is available. It seems that oxytocin does not amplify the inclination to cooperate in itself. Subjects given oxytocin rather anticipate that defective behavior would provoke negative emotions of others, which would ultimately trigger the punishment of their defective behavior (Fehr & Gächter, 2002). And indeed, the social norm of cooperation is defended aggressively in the prisoners' dilemma game due to our oxytocin treatment. Accordingly, our results suggest that oxytocin has no influence on general pro-social attitudes, but rather induces a negative emotional reaction to norm-violating behavior. As social norms constitute a vital element in the maintenance of social groups, our findings contribute to recent evidence that emphasizes the important role of oxytocin in group cohesion (De Dreu et al., 2010; De Dreu et al., 2011).

The second chapter, which is joint work with Andrea Jobst, Martin Kocher, Fabian Loy and Norbert Müller, links oxytocin to the adherence to a widely accepted norm, i.e. honesty. In particular, we employ a coin tossing task in which the experimenter is not able to detect lies on an individual level, but to infer dishonesty on an aggregate level in an unobtrusive way. We test the role of oxytocin in dishonesty in two treatments. In the first treatment, we elicit subjects' general inclination to lie for their own financial benefit in a situation with no negative externalities on others. In the second treatment, we introduce a competitive environment with negative externalities on competitors, as dishonesty pays only for the one who lies more than another randomly matched participant. In particular, we want to test whether an exogenous alteration of brain oxytocin levels modulates the inclination to break (or bend) a widely accepted norm, i.e. honesty, in a competitive task compared to a non-competitive situation.

Despite the fact that competition and the institution of a (perfect) market with no externalities have positive effects on social welfare, it has been shown that keen competition in markets may also pressure companies to engage in unethical behavior. The argument of immorally behaving competitors is a commonly used justification to consider own unethical behavior as a necessity to keep up with competitors. We argue that this may even lead to the view that as long as all others do it, it can be considered as less harmful or even appropriate (Falk & Szech, 2013; Gino, Ayal, & Ariely, 2009).

Our results provide evidence that oxytocin had a detrimental effect on honesty in a competitive environment, as lying rates significantly increased compared to a non-competitive environment. This effect was not present in the placebo group. Furthermore, we provide evidence that the effect is mainly rooted in a higher inclination to conform to others' unethical behavior. We contribute to previous empirical findings that demonstrated the detrimental effects of competition on morals (Belot & Schröder, 2013; Falk & Szech, 2013; Faravelli, Friesen, & Gangadharan, 2015; Schwieren & Weichselbaumer, 2010). Particularly, our work extends previous findings by providing evidence that in-group conformity plays a crucial role in competitive environments, as unethical behavior is perceived as more permissible during competition. Our results suggest that subjects who are motivated to preserve a positive self-image are more prone to such justifications.

Chapter 3, which is joint work with Sandra Dehning, Andrea Jobst, Martin Kocher, Norbert Müller and Peter Zill, examines the effect of oxytocin on the inclination of individuals to form groups. This work is primarily motivated by the growing body of studies that document the important role of oxytocin in modulating group behavior, since the neuropeptide positively influences in-group conformity (Stallen, De Dreu, Shalvi, Smidts, & Sanfey, 2012), in-group favoritism (De Dreu et al., 2011) and in-group cooperation (De Dreu et al., 2010). Despite these positive effects on group cohesion, little or nothing is known about the role of the oxytocinergic system in the endogenous formation of groups. In particular, this work examines the effects of oxytocin on the inclination to form groups and to engage in collaborative decision-making, despite the risk of subordinating to inferior decisions of others. In our experimental design, subjects are asked to choose solving an incentive compatible task either alone or as a member of a team with three (randomly chosen) subjects. Particularly, subjects have to indicate whether they want to play the *p*-Beauty contest game – a commonly used tool to measure strategic sophistication – in teams or as individuals. That is, the *p*-Beauty contest game constitutes an ideal tool for studying the performance of teams and individuals, as it allows to measure the steps of reasoning that a decision-maker actually applies to solve this task (Camerer, Ho, & Chong, 2004; Nagel, 1995).

Contrary to our initial hypothesis, our data provides evidence that an exogenous alteration of brain oxytocin levels causes a significant adverse effect on group formation, as subjects given oxytocin exhibit a significantly lower propensity to opt for team decision-making compared to placebo. This finding questions standard hypotheses regarding the role of oxytocin in modulating group behavior (Bethlehem et al., 2014; MacDonald & MacDonald, 2010).

Moreover, we find that this adverse effect is most pronounced for subjects with relatively high strategic sophistication, as they show a significantly lower propensity to opt for team decision-making due to our treatment. Our analysis suggests that this adverse effect of oxytocin is mainly driven by a shift in beliefs regarding the success of teams in this task. Whereas subjects in the placebo group significantly over-estimate the chances of teams, subjects given oxytocin have more realistic (i.e. downward revised) beliefs about the winning frequency of teams. Consequently, this work questions standard hypotheses regarding the role of oxytocin in modulating group behavior and group cohesion.

Furthermore, our data provides initial evidence that oxytocin modulates strategic thinking, as subjects given oxytocin showed a significantly enhanced performance in the strategic task. However, we are not able to draw general inferences regarding the positive effects of oxytocin on strategic thinking, since our experimental design includes a self-selection process and additionally diverse decision-makers (teams vs. individuals). Accordingly, additional research on this issue is indispensable to infer general conclusions on the effects of oxytocin on strategic thinking, which might ultimately lead to a new hypothesis that is able to explain all well-documented effects of oxytocin in a unified approach.

Chapter 4 contributes to the scarce literature regarding the modulating role of hormones on strategic thinking. Grounded on evidence that oxytocin enhances cognitive empathy, i.e. the ability to put oneself emotionally in the shoes of another person (Domes et al., 2007), this work examines the effects of an exogenous alteration of brain oxytocin levels on strategic perspective taking. In particular, due to its positive effects on the emotional part of mentalizing, I test the hypothesis that oxytocin has a positive effect on strategic thinking in a competitive task.

To test this hypothesis, I employ the *p*-Beauty contest game as a main paradigm, since it constitutes an ideal tool to measure strategic sophistication by eliciting the exact applied steps of reasoning to solve this task. In the experiment, participants are randomly assigned to units of four and are asked to play the *p*-Beauty contest game for four rounds with feedback after each round. The winner in each round receives a real monetary prize (i.e. 10 € per round), or a split of the prize in case of a tie. Furthermore, I elicit subjects' cognitive ability as well as their overconfidence, since previous studies showed that a rational guess (i.e. best response) in this game depends on one's own cognitive ability (Burnham, Cesarini, Johannesson, Lichtenstein, & Wallace, 2009; Carpenter, Graham, & Wolf, 2013), and additionally, on her beliefs about the average sophistication of her opponents (Agranov, Potamites, Schotter, & Tergiman, 2012; Camerer & Lovallo, 1999). That is, by excluding the possibility that oxytocin might influence one's general cognitive ability or one's beliefs about the others' sophistication, it is possible to trace back the effect of oxytocin to a modulation in strategic sophistication. Therefore, I elicit general cognitive ability with a standard non-verbal assessment of general intelligence, the Raven's advanced progressive matrices task (Raven, 1965), and additionally subjects' ex ante beliefs about their co-players performance in the *p*-Beauty contest game. Accordingly, if neither general cognitive ability nor the assessment of others' intelligence is altered by oxytocin, then any change in the behavior of subjects should be caused by a modulation of strategic sophistication due to the oxytocin treatment.

The results indicate that oxytocin has a positive effect on strategic reasoning, which is most pronounced for subjects with relatively low general cognitive ability. These findings contribute to previous studies that demonstrated the important role of oxytocinergic systems in (emotional) perspective taking (Domes et al., 2007). Moreover, I employ the cognitive hierarchy model to provide a computational measure of the exact applied steps in strategic thinking to solve this task (Camerer et al., 2004). According to the results, oxytocin almost double the utilized level of reasoning compared to placebo. Moreover, participants in the oxytocin condition require 45% more time to decide in the *p*-Beauty contest game, whereas response times in a neutral task show no difference. This indicates that more steps of reasoning require more time to solve this task. This is in line with the notion that strategic reasoning, like reasoning in general, is a slow and effortful process rather than a fast and impulsive one (Frederick, 2005; Metcalfe & Mischel, 1999). This view seems to be supported by previous evidence showing that brain areas related to higher cognitive functions are involved in strategic reasoning and mentalizing (Amodio & Frith, 2006; Carrington & Bailey, 2009; Coricelli & Nagel, 2009), and longer decision times are associated with greater performance in the *p*-Beauty contest game (Chong, Camerer, & Ho, 2005; Kocher & Sutter, 2006; Rubinstein, 2007).

This work provides novel evidence regarding the neurological basis of economic decision-making by showing that strategic reasoning (or strategic mentalizing) is partially encoded in special neural substrates related to the oxytocinergic system (Amodio & Frith, 2006; Coricelli & Nagel, 2009). As social beings, humans are constantly forced to infer the strategies of others which is a prerequisite for coordination, cooperation (De Dreu et al., 2010) and trust (Kosfeld et al., 2005; Zak et al., 2005). In view of that, it is not surprising that a dysfunction of the oxytocinergic system is related to several disorders characterized by impairments in social interaction and mentalizing abilities, including particular forms of autism spectrum disorder (Andari et al., 2010; Guastella et al., 2010; Jacob et al., 2007; Modahl et al., 1998) and schizophrenia (Goldman, Marlow-O'Connor, Torres, & Carter, 2008; Pedersen et al., 2011). In light of this evidence, the oxytocinergic system may constitute a necessity to enable humans to engage in strategic thinking or strategic mentalizing, which is indispensable to sustain larger societies with a tremendous set of social interactions that require a huge amount of cognitive capacity (Dunbar, 1998; Dunbar & Shultz, 2007).

Each of the four chapters in this dissertation can be read independently, since all chapters are self-contained and include their own introductions and appendices. The respective appendices contain detailed information of the experimental protocols, instructions as well as supplementary analyses, figures and tables.

Chapter 1

Oxytocin promotes altruistic punishment*

1.1 Introduction

A remarkable element of human societies is the ability to cooperate and to enforce cooperation, even among genetically unrelated strangers in one-shot interactions (Bernhard, Fischbacher, & Fehr, 2006; Falk, Fehr, & Fischbacher, 2005; Fehr & Gächter, 2002). Altruistic punishment has been identified as probably the most important mechanism to enforce cooperation (Fehr & Gächter, 2000; Ostrom, Walker, & Gardner, 1992; Yamagishi, 1986). However, evidence on the neurological foundations of punishment that enforces the social norm of cooperation is still scarce (De Quervain et al., 2004; Egas & Riedl, 2008; Fehr & Fischbacher, 2004). Since altruistic punishment is costly to the punisher and to the punished, it requires either a pro-social attitude or a negative emotional reaction to norm-violating behavior or both (Fehr & Gächter, 2002; Hopfensitz & Reuben, 2009).

In this work, we linked oxytocin, an endogenous neuropeptide produced in the hypothalamus, to the inclination of using altruistic punishment. Oxytocin is well-known for promoting pro-social behavior (MacDonald & MacDonald, 2010; Zak et al., 2007), to enhance social bonds (Ditzen et al., 2009) and to foster trust (Baumgartner, Heinrichs, Vonlanthen, Fischbacher, & Fehr, 2008; Kosfeld et al., 2005; Zak et al., 2005). Cooperation in a one-shot social dilemma requires trust in the cooperation of others as well as a pro-social attitude. Hence, one would expect oxytocin to promote cooperation, even without the option to enforce the norm of cooperation. However, previous studies of the neuropeptide provided

* This chapter is based on joint work with Nadja Furtner, Andrea Jobst, Bianca Kern, Martin G. Kocher und Norbert Müller.

results on more mixed effects on cooperation, as oxytocin promotes cooperation with in-group members, but not with a competing out-group (De Dreu et al., 2010). Moreover, exogenously applied oxytocin increases the sensitivity to social information (Bartz et al., 2011; Declerck et al., 2010; Declerck et al., 2013).

It has even been shown that oxytocin is able to increase negative social emotions in situations perceived as unfair (Shamay-Tsoory et al., 2009). We argue that this may also induce altruistic punishment, since negative emotions can serve as a proximate mechanism in norm enforcing behavior (Fehr & Gächter, 2002). That is, when a punishment option is available, punishers incur the cost of enforcing the norm without receiving any material benefits. But cooperation can still be enforced if enough altruistic punishers are present (Fehr & Schmidt, 1999), and if their sanctioning behavior is triggered through negative social emotions as a consequence of the payoff difference between being a free rider (defector) and being a co-operator. However empirical evidence on the influence of oxytocin on the enforcement of a social norm is still missing. Based on findings on negative social emotions, we hypothesize that oxytocin enhances the enforcement of social norms.

1.2 Experimental design

We used a double-blind study design in which participants intra-nasally self-administered either oxytocin ($N = 72$) or a placebo ($N = 72$) before interacting in a one-shot, simultaneous prisoners' dilemma game (following Falk et al., 2005). The one-shot prisoners' dilemma game was implemented without a punishment (sanctioning) opportunity (henceforth, PDGX) and with a punishment opportunity (henceforth, PDGS). It involved real monetary stakes (MU for "monetary units" in the following; converted to euro amounts). Participants were randomly and anonymously matched to groups of three. Game PDGX consisted of a single choice to cooperate or to defect; payoffs of this social dilemma are displayed in Table 1.1. Accordingly, individual payoff maximization predicts defecting of all players, whereas collectively, cooperation would be the social optimal outcome.

	Both other players defect	**One of the other two players cooperates**	**Both other players cooperate**
Player *i* defects	20	32	44
Player *i* cooperates	12	24	36

Table 1.1 | Payoffs to player *i* in PDGX and the first stage of PDGS. Participants were randomly and anonymously assigned to groups of three. They decided simultaneously whether to cooperate or to defect. For example, if player *i* decided to defect whereas both other group members decided to cooperate, player *i* would earn an income of 44 monetary units (MU) in PDGX and/or in the first stage of PDGS. If instead player *i* decided to cooperate, he would earn an income of 36 MU.

Game PDGS consisted of two stages: in the first stage participants played a three-person prisoners' dilemma game equivalent to PDGX. In the second stage of the PDGS participants had the opportunity to sanction. They were first informed about the decisions (cooperate/defect) of the other two group members in stage 1. Then, they could assign up to a maximum of ten punishment points to each of the other two group members, neutrally denoted "deduction points" in the experimental instructions. Each assigned punishment point cost the punishing participant 1 MU, whereas the recipient's payoff was deducted by 3 MU per point. The final individual payoff from PDGS was the income from the first stage, minus three times the sum of received punishment points, minus the sum of assigned punishment points (see *Methods*). As punishment was costly, the dominant strategy for selfish decision makers was not to punish (see appendix A1 for details). All design details were common knowledge among experimental participants, and control questions at the beginning of the experiment made sure that they fully understood them.

1.3 Results

1.3.1 The effect of oxytocin on cooperative behavior with and without punishment

In game PDGX, we observed no significant treatment effect of oxytocin (see Fig. 1.1). However, the effect went in the expected direction: with oxytocin, the average level of cooperation is at 30.6%, compared to 20.8% in the placebo group ($\chi^2 = 1.7823$; $df = 1$; $P = 0.182$, two-sided). In game PDGS, with punishment, cooperation levels were significantly higher—a result that has been established multiple times in the literature (Chaudhuri, 2011; Falk et al., 2005; Fehr & Gächter, 2000, 2002). However, there was an additional effect of oxytocin that is significant. Cooperation levels reached 52.8% in the placebo group and even 69.4% in the oxytocin group (χ^2 = 4.2078; *df* = 1; *P* = 0.04, two-sided).

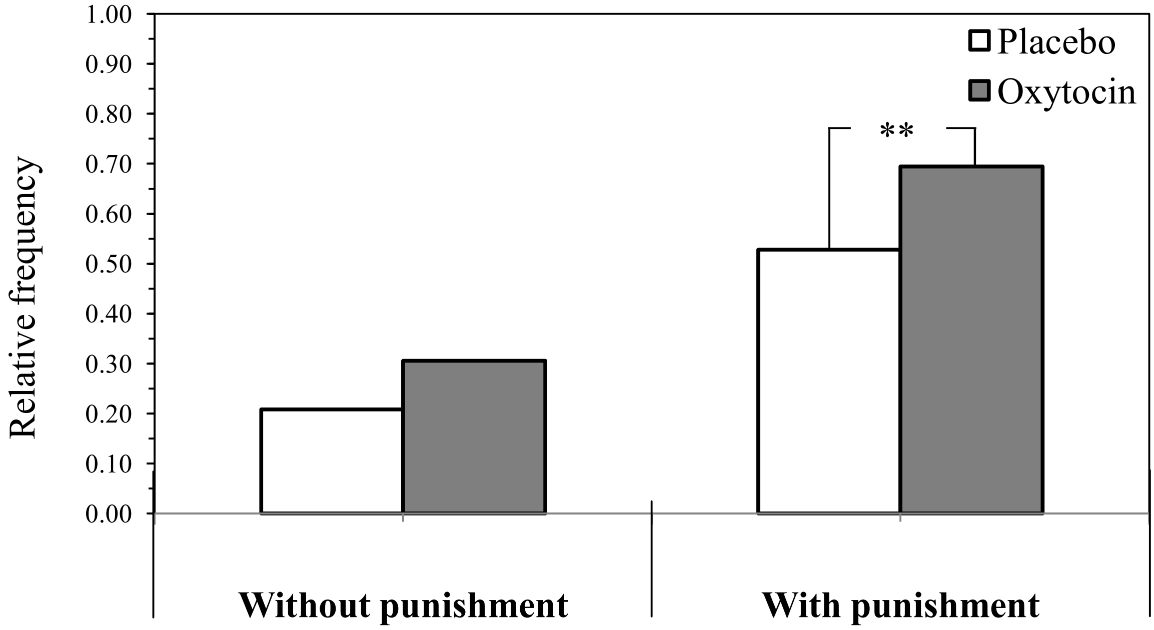


Fig. 1.1 | Cooperation with and without a punishment option. Relative frequency of participants who decided to cooperate with and without a punishment option in the oxytocin (white bars) and placebo (gray bars) groups (*N* = 144). Without a punishment option, participants in the oxytocin group and the placebo group exhibited no significant difference in cooperative behavior. In the presence of a punishment option, participants in the oxytocin group had a significantly higher inclination to cooperate.

1.3.2 The effect of oxytocin on punishment behavior

We now turn to punishment behavior. Our data showed that punishment imposed on defectors was considerably higher under the influence of the neuropeptide (see Fig. 1.2). In the oxytocin group, defectors received, on average, 4.3 punishment points, which were almost twice the amount of the 2.2 punishment points imposed on defectors in the placebo group (Mann-Whitney *U*-test; $z = -2.334$, $P = 0.0196$, two-sided, $N = 22$ in the oxytocin and $N = 34$ in the placebo group). When cases of zero punishment are excluded, imposed sanctions were about 34% higher (6.3 vs. 4.7 points) in the oxytocin group than in the placebo group (Mann-Whitney *U*-test; $z = -2.139$, $P = 0.032$, two-sided, $N = 15$ in the oxytocin and $N = 16$ in the placebo group). We conclude that not only the inclination to punish was enhanced with oxytocin, but also the degree of punishment. In contrast, punishment of cooperators was virtually absent (see Fig. 1.4 in appendix A2), without any statistical difference between the oxytocin group and the control group (Mann-Whitney U-test; z = –0.578, P = 0.5630, two-sided, $N = 50$ in the oxytocin and $N = 38$ in the placebo group).

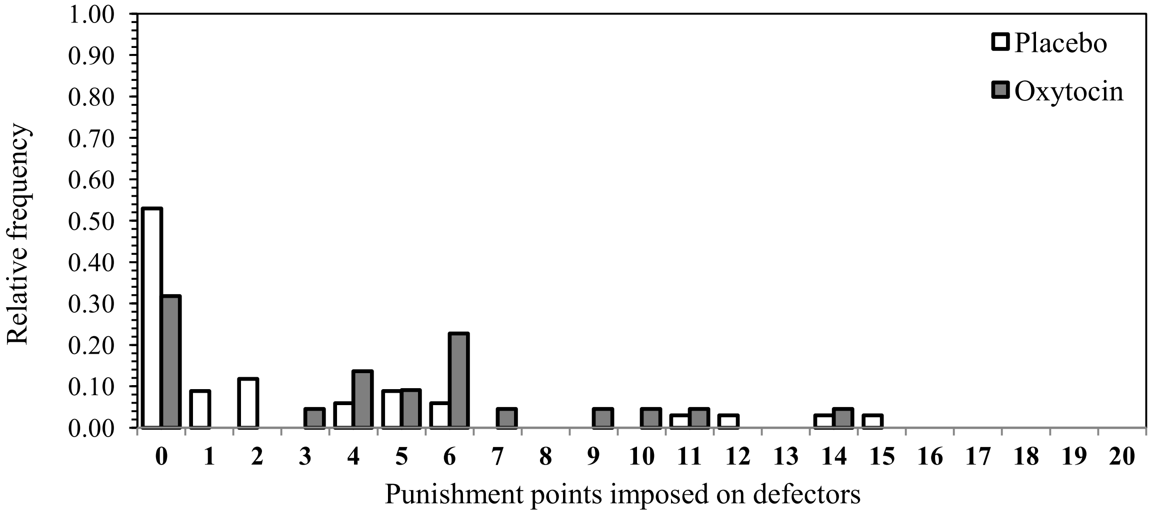


Fig. 1.2 | Punishment points imposed on defectors. Relative frequency of assigned punishment points in the oxytocin (gray bars) and placebo (white bars) groups (N = 144). Punishment points imposed on defectors were significantly higher in the oxytocin group than in the placebo group.

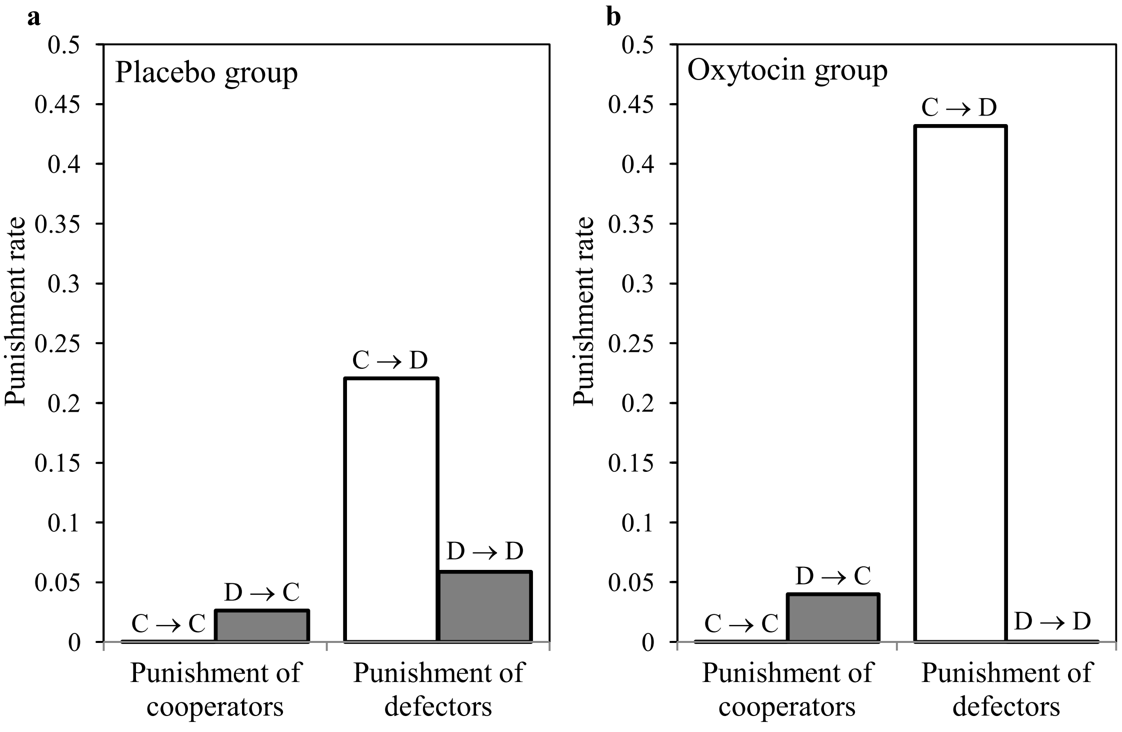


Fig. 1.3 | Punishment rate by punished types. The bar C→C represents the proportion of all cooperators punished by other cooperators; analogously, C→D is the proportion of all defectors punished by cooperators; and so on. (a) In the placebo group about 22% (15 out of 68) of all defectors were punished by cooperators, whereas only 6% (4 out 68) were punished by other defectors. Punishment of cooperators (2 out of 76) played only a minor role. (b) In the oxytocin group 43% (19 out of 44) of all defectors were punished by cooperators. A few defectors punished cooperators (4 out of 100).

Fig. 1.3 displays the direction of all punishment activity and confirms that, in any case, the vast majority of punishment incidents went from cooperators to defectors. Hence, punishment was clearly fairness-induced in the treatment group and in the control group. However, there was a clear difference in the inclination to punish, as already indicated in Fig. 1.2. In the placebo group, roughly 22% (15 out of 68) of all defectors were punished by cooperators. In contrast, about 43% (19 out of 44) of defectors in the oxytocin group were punished, which is significantly higher than the punishment rate in the placebo group (χ^2 = 5.6379; *df* = 1; *P* = 0.018, two-sided).

One important question is whether oxytocin might have triggered a stronger inclination to punish others, regardless of the violation of norms, for instance, as a consequence of a spiteful motive or an undirected emotional arousal. Such reasoning would imply that the punishment behavior of defectors against both cooperators and other defectors would differ between the oxytocin group and the placebo group. From eyeballing Fig. 1.3, it is obvious that this was not the case. Punishment incidents by defectors were not more frequent in the oxytocin group than in the placebo group (4 out of 44 in the oxytocin group vs. 6 out of 68 in the placebo group; Fisher's exact-test; *P* = 1, two-sided). Given the small numbers, the directions of punishment were obviously not different between defectors in the oxytocin group and in the placebo group (see appendix A3 for a supplementary regression analysis of cooperation and punishment decisions).

Consequently, oxytocin had a weak, but insignificant effect on the inclination to cooperate in our setup. However, it significantly promoted both the likelihood and the intensity of punishment compared to the control group. However, the structure of punishment remained unaltered. That is, both, in the control group and in the oxytocin group, the vast majority of punishment incidents were imposed by cooperators on defectors, which enforced a social norm of cooperation. Hence, oxytocin increased neither spiteful punishment nor the general inclination to punish others, but rather enhanced the propensity to enforce the norm of cooperation.

1.3.3 Oxytocin and negative emotions

A final question is whether our results on oxytocin promoting punishment are consistent with emotional responses. We hypothesize that the effect of oxytocin on punishment behavior might be connected to the modulation of negative emotions, triggered by the violation of a social norm. Specifically, negative emotions related to the violation of fairness norms, such as anger and disappointment, would be intensified under the influence of oxytocin and thus trigger punishing behavior (Shamay-Tsoory et al., 2009). In order to test this hypothesis, participants were asked to report the intensity of several positive and negative emotions on a seven-point Likert scale immediately after the allocation of their punishment points and – in order to control for baseline effects – at the beginning of the experiment, subsequent to the substance administration. As usual in survey questions on emotions, a list of related and probably unrelated emotions was presented in order to avoid experimenter demand effects in elicitation. (A focus here was on anger and disappointment; details on the rating of other emotions can be found in Table 1.4 in the appendix A4). In fact, participants given oxytocin reported higher levels of anger towards defectors (3.07 vs. 1.65; Mann-Whitney U-test; $z = -2.701$, $P = 0.007$, two-sided, controlled for baseline emotions) and higher levels of disappointment (2.95 vs. 2.20; Mann-Whitney U-test; $z = -1.800$, $P = 0.0718$, two-sided, controlled for baseline emotions). Additionally, according to a survey question, participants in the oxytocin group perceived the behavior of defecting group members as significantly less fair than participants in the placebo group (2.93 vs. 3.39; Mann-Whitney U-test; $z = 1.860$, $P = 0.0629$, two-sided). This indicates that participants given oxytocin were emotionally more prone to anger in the face of violated social norms and suggests a significant impact of oxytocin on the affective system in the brain involved in emotional reactions. It is moreover consistent with the finding that it is also emotions in addition to rational considerations that work as a proximate mechanism to induce norm enforcing behavior (De Quervain et al., 2004; Falk et al., 2005; Fehr & Gächter, 2000; Hopfensitz & Reuben, 2009).

1.4 Conclusion

Social norms are ubiquitous in human society, and many of the norms require enforcement to be stabilized. Oxytocin does not only seem to make norm adherence more likely through a positive impact on pro-social behavior in certain but not all settings. An equally important role of oxytocin has been investigated and substantiated in this study: oxytocin promotes the likelihood and the intensity of norm enforcement activities. The channel through which it operates seems an augmented arousal of negative emotions following a norm violation that induces pro-social decision makers. Our findings extend existing studies that have indicated evidence for oxytocin to stimulate negative emotions (Shamay-Tsoory et al., 2009) or defensive aggression. Former studies have shown that oxytocin promotes defensive aggression towards members of a competing out-group, but at the same time increases cooperation and trust towards in-group members (De Dreu et al., 2010; van IJzendoorn & Bakermans-Kranenburg, 2012). In our case, it was the social norm that tied the group together and was defended aggressively in the prisoners' dilemma game. Our results call for future work on the connection of oxytocin, social norms and norm-enforcement behavior, with the ultimate goal to look at potential clinical implications of oxytocin in disorders characterized by a large disparity between behavior and prevailing social norms.

1.5 Methods

1.5.1 Participants

A total of 144 male participants from different universities in Munich participated in the experiment. Participants were between 18 and 33 years old (mean age ± s.d., 23.7 ± 3.1 years) and were recruited by ORSEE (Greiner, 2004). Exclusion criteria for participation were significant medical or psychiatric disorder, medication, smoking more than 15 cigarettes per day, and drug or alcohol abuse. Participants were instructed to abstain from alcohol, smoking and caffeine for 24 h, and from food for 1 h before the experiment. At the time of recruitment participants were informed that the experiment is about the influence of oxytocin on (economic) decision making. The study was approved by the ethics committee of the medical department at the University of Munich. All participants provided written informed consent before participation.

1.5.2 Substance administration

We used a double-blind, placebo-controlled study design to compare behavior in a group of participants that received a self-administered, single intranasal dose of 24 IU oxytocin (Syntocinon-Spray, Sigma-Tau; 3 puffs per nostril, each with 4 IU of oxytocin) with that of participants in a control group that received a placebo. Participants were randomly assigned to the two groups and received oxytocin or a placebo 50 minutes before the relevant part of the experiment started. In one experimental session, either oxytocin or a placebo was administered to all participants. The placebo contained all inactive ingredients except for the neuropeptide (Born et al., 2002; Gossen et al., 2012).

1.5.3 Behavioral experiment and questionnaires

Participants received neutrally written instructions (see *Instructions* in the appendix A5), which we read aloud to make them common knowledge. They knew that the experiment consisted of several parts which were independent of each other. Instructions for one part only followed after the part before was accomplished. Decisions were anonymously made on PC screens in separated cubicles. Fifty minutes after substance administration, we started with the incentivized experiment, which consisted of a three-person prisoners' dilemma game without a sanctioning opportunity (PDGX) and a three-person prisoners' dilemma game with a sanctioning opportunity (PDGS). In both treatments, PDGX and PDGS, we ensured that participants understood the payoff structure by control questions. In both treatments, participants were randomly and anonymously assigned to groups of three. They knew that they were not matched with the same persons in the two treatments or another part of the experiment (perfect stranger matching). We conducted one period of PDGX and PDGS each.

In PDGX participants had to decide simultaneously in groups of three whether to cooperate or defect. The incentives of the PDGX game are outlined in Table 1.1. Feedback on the other group members' decisions and on the payoff from PDGX was only given at the end of the entire experiment.

The subsequent PDGS game consisted of two pre-announced stages: In the first stage participants played again the simultaneous three-person prisoners' dilemma game. In the second stage participants received feedback on the other two group members' decisions in the first stage and then had the opportunity to reduce their individual payoff by assigning deduction points. One deduction point cost the punisher 1 MU and the punished person 3 MU. Each participant could assign up to 10 punishment points to each of the other two group members. To avoid that participants earn negative incomes in PDGS without being able to influence them, a participant's total income from the two stages could not be reduced by the other two group members below 0 MU.

To measure emotional states of participants throughout the experiment and especially during the norm enforcement decision in PDGS, participants rated 8 emotions on a 7-point Likert scale. Participants had to assess positive and negative emotions, namely anger, gratefulness, guilt, joy, irritation, shame, surprise and disappointment (Hopfensitz & Reuben, 2009). The rating was conducted immediately after substance administration and in PDGS. The second assessment took place immediately before the norm enforcement decision in PDGS, and we additionally asked participants to rate the fairness of the other group members' decisions in the first stage. Participants received an additional 20 MU for the completion of all questionnaires. The whole experiment lasted up to 120 minutes and was programmed using z-Tree software (Fischbacher, 2007). Each MU earned in the experiment was worth 0.2 €. The final payoff consisted of the sum of incomes from the different parts and a flat fee of 4 € for participation.

Appendix A

A1. Proofs and explanations of the predictions

The Prisoners Dilemma Game without Punishment: All subjects are randomly assigned to groups of three and given an implicit initial endowment of 20 monetary units (which corresponds to the amount of monetary units if everyone chooses not to cooperate). Then, they have to decide if they want to cooperate and invest their whole endowment into a public good or to defect and invest nothing. The investment choice g hence is binary with $g \in \{0, 20\}$ and the income π_i of subject i is given by

$$\pi_i = 20 - g_i + 0.6 \sum_{j=1}^{3} g_j,$$

where g_i denotes the contribution of subject i and the sum of g_j denotes the contributions of all three group members to the public good. The marginal per capita return (MPCR) from investing in the public good is 0.6. From a collective perspective, it would hence be optimal if all subjects contribute their entire endowment to the public good, as $n * MPCR = 1.8 > 1$. From an individual perspective however, free-riding is a dominant strategy for every subject, as it is always profit maximizing:

$$\frac{\partial \pi_i}{\partial g_i} < 0.$$

Consequently, this leads to the prediction (or Nash-Equilibrium) that all subjects will defect and not invest into the public good, i.e. $g_i = g_j = 0$ and $\pi_i = \pi_j = 20$.

The Prisoners Dilemma Game with Punishment: All subjects are again randomly assigned to groups of three. In the first part they have to decide if they want to cooperate and invest into the public good or defect and invest nothing. The investment choice g is hence again binary with $g \in \{0, 20\}$. In the second part, subjects additionally have to decide if they want to costly reduce their group members profits and invest in deduction (punishment) points p, where each deduction point costs one monetary unit and reduces the payoff of the punished person by three monetary units. While the deduction points can reduce the punished person's payoff only to a minimum of zero, subjects still have to bear the costs for allocated deduction points in any case.

The income π_i of subject i is hence given by

$$\pi_i = \max\left[0, 20 - g_i + 0.6\sum_j g_j - 3\sum_{j\neq i} p_{ji}\right] - \sum_{j\neq i} p_{ij},$$

where g_i denotes the contribution of subject i, the sum of g_j denotes the contributions of all three group members to the public good, p_{ij} is the amount of deduction points subject i assigns to subject j with $i \neq j$, and p_{ji} is the amount of deduction points subject i receives by subject j.

Since there are no future interactions, punishment of other group members has no future material gains but incurs immediate losses to the punishing person. So profit-maximizing subjects should never invest into sanctioning of group members, as the costs they have to bear always reduce their payoffs, i.e. $\frac{\partial \pi_i}{\partial p_{ij}} < 0$.

However, if no one allocates any punishment points, this game reduces to the case with no punishment option. Thus, the marginal per capita return (MPCR) from investing in the public good is again 0.6. And again, from a collective perspective, it would be optimal if all subjects would invest into the public good, as $n * MPCR = 1.8 > 1$.

However, from an individual perspective, free-riding is still a dominant strategy for every subject, as $\frac{\partial \pi_i}{\partial g_i} < 0$.

Consequently, this leads to the prediction that subjects will never invest into the public good and never invest into deduction points to punish other group members, i.e. $g_i = g_j = 0$, $p_{ij} = p_{ji} = 0$ and $\pi_i = \pi_j = 20$.

A2. Punishment points imposed on cooperators

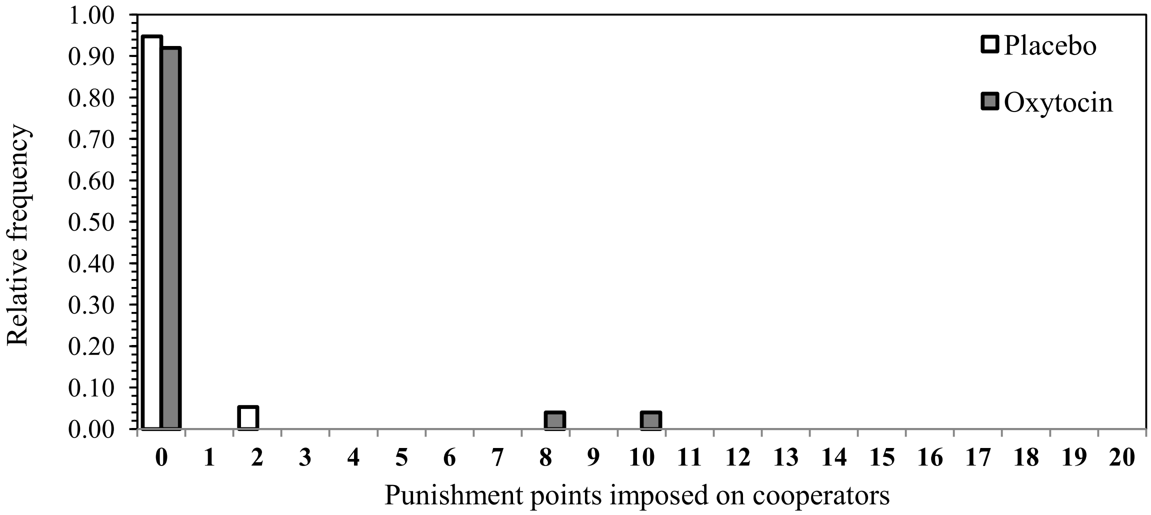


Fig. 1.4 | Punishment points imposed on cooperators. Relative frequency of assigned punishment points in the oxytocin (gray bars) and placebo (white bars) groups on cooperators (N = 88). Punishment of cooperators was almost not existent, and additionally, subjects in the oxytocin group and the placebo group showed statistically identical inclination to punish cooperators (Mann-Whitney U-test; z = –0.578, P = 0.5630, two-sided).

A3. Additional data analysis of punishment and cooperation

Punishment behavior controlled for overall cooperation levels:

We conducted an additional regression analysis of the effect of oxytocin on punishment decisions in order to control for other variables that might influence this behavior. Particularly, we were interested in the effect of oxytocin on punishment behavior if controlled for the overall cooperation level within a group. The overall cooperation level might have influenced punishment behavior, since punishment could be different if only one other group member defected compared to the case where both other group members defected. Therefore, it is necessary to show that the described effect of oxytocin on punishment was still present even if controlled for the overall cooperation level. In the regression analysis of Table 1.2, the dependent variable is the amount of received punishment points of subject *i* assigned by another group member *j*.

	Dep. Variable: ***Received deduction points of participant i***	
Ind. Variables	(1)	(2)
Others' cooperation level	.542*	.260
Negative deviation of i	3.161***	2.181***
Positive deviation of i	.427	-.169
Negative deviation of i (OT=1)		2.024*
Positive deviation of i (OT=1)		.690
OT		-.043
Constant	-.102	.319
R^2	0.217	.243

Table 1.2 | Influences on received deduction points. Note that ***, **, * denote significance at the 1%, 5%, 10% level. The amount of received punishment points of *i* is the dependent variable in a OLS regression with $N = 288$ observations and clustered standard errors on participant level. A Tobit regression provides similar results. An post-estimationWald-test reveals that the coefficients of OT+NegativeDeviation(OT=1) is significantly different from 0 (p=0.0960).

There are 4 independent variables, "other's cooperation level", "negative deviation", "positive deviation" and "OT" (=Oxytocin). The "others' cooperation level" is defined as 0 if both other group members (other than *i*) defected, 1 if one of them cooperated and 2 if both other group members chose to cooperate. The variable “negative deviation” is 1 if subject *i* defected and player *j* cooperated and 0 otherwise. The variable “positive deviation” is 1 if subject *i* cooperated and player *j* defected and 0 otherwise. The variable Oxytocin indicates the vero (=1) and the placebo group (=0).

Regression (1): The first regression shows the effect of a positive or negative deviation on the received punishment points controlled for the others’ cooperation level. Obviously, only a negative deviation resulted in a significantly higher punishment, whereas the other’s cooperation level was only weakly significant.

Regression (2): The second regression is extended by the variables “oxytocin” and an interaction effect between “oxytocin” and a “positive or negative deviation”. Interestingly, the overall cooperation level seems to have no significant effect if oxytocin is included in the regression. That is regardless of the overall cooperation level, only defectors were punished. Furthermore, oxytocin seems to have a highly conditional effect on punishment, since only the interaction variable of oxytocin and a negative deviation is significant. Therefore, oxytocin itself had no general increasing effect of punishment. It rather had a very conditional and specific effect. That is, subjects given oxytocin punished a negative deviation significantly harsher compared to placebo, but not in the case of a positive deviation. As a result, oxytocin had an additional effect on punishment of defectors that led to an almost doubled amount of punishment points imposed on defectors. This is in line with the described effect in the main article.

Influences on the inclination to cooperate:

The following Table 1.3 shows a Probit regression of a subject’s choice to cooperate on the variables oxytocin, the possibility to punish and the interaction effect of both. The mere presence of a punishment option significantly increased cooperation levels. This is in line with the established insight that cooperation can be stabilized by the mere presence of a sanctioning mechanism (Falk et al., 2005; Fehr & Gächter, 2002). However, oxytocin itself had no significant effect on the inclination to cooperate, whereas the interaction effect of both variables is significant. This indicates that oxytocin had no general effect on cooperation, but rather a modulating effect. That is, subjects given oxytocin reacted even stronger on the presence of a punishment option, resulting in an even higher cooperation level.

Dep. Variable: ***Choice to cooperate*** (*defect = 0; cooperate = 1*)	
Ind. Variables	
Punishment Option	.340***
Oxytocin	.119
Oxytocin× Punishment Option	.486***
Pseudo R^2	.1108

Table 1.3 | Influences on cooperation choices. Note that ***, **, * denote significance at the 1%, 5%, 10% level. The choice to cooperate is the dependent variable in a Probit regression with clustered standard errors on participant level. Positive coefficients indicate a positive marginal effect on probability to cooperate.

A4. Emotions towards defectors

Emotions	(Placebo)		(Oxytocin)
Anger	1.64 (.311)	––––***––––	3.06 (.390)
Gratefulness	-1.98 (.202)		-1.56 (.284)
Guilt	-0.42 (.193)		-0.40 (.166)
Joy	-2.14 (.254)		-2.50 (.304)
Irritation	0.35 (.233)	––––***––––	1.68 (.269)
Shame	-0.26 (.139)		-0.15 (.197)
Surprise	0.63 (.248)	––––**––––	1.52 (.338)
Disappointment	2.20 (.270)	––––*––––	2.95 (.372)
Perception of co-players actions	(Placebo)		(Oxytocin)
Fairness	3.39 (.202)	––––*––––	2.93 (.315)

Table 1.4 | The effect of oxytocin on emotions towards defectors. The table shows eight different general emotions towards defectors compared between the placebo and the oxytocin group. The emotional rating towards a defecting group member was subtracted by baseline emotions that were elicited immediately after substance administration. Note that ***, **, * denote significance at the 1%, 5%, 10% level and standard errors are in parentheses (Mann-Whitney U-tests; two-sided).

A5. Instructions (translated from German)

Welcome to the experiment and thank you for your participation!

From now on, please do not talk to other participants and please switch off your phone!

General Procedure

The purpose of this experiment is to study decision making. You can earn money during this experiment which will be paid out to you in cash at the end of the experiment.
During the experiment you and the other participants will make decisions. Your own decisions as well as those of the other participants will affect your payment according to the rules explained in the following.
The duration of the experiment is **120 minutes**. If you should have any questions, please raise your hand. One of the experimenters will come to you and answer your questions in private.
During the experiment we do not speak of Euros, but of **Guilders**. Thus, your earnings during the experiment will be calculated in Guilders**.** At the end of the experiment your earnings in Guilders will be converted into Euros according to the following exchange rate:

5 Guilders = 1 Euro
1 Guilder = 20 Eurocent

For your punctual arrival you additionally receive 20 Guilders. At the end of the experiment you will be paid your total earnings from the experiment in private and in cash.
While you make your decisions a clock will count down at the upper right corner of the computer screen. This countdown serves as an orientation for how much time you should need to take the decision. However, usually you can take more time to make your decision. If it is not possible to take more time than indicated, we will inform you in advance.

Anonymity
Your decisions over the course of the experiment are anonymous. You will not receive personal information about the other participants of the experiment. All data elicited will be encoded and will be treated confidentially. At the end of the experiment you will sign a confirmation which says that you have received the payment. This confirmation only serves accounting purposes for our sponsor. Our sponsor will not receive any individual data from the experiment either.

Utilities
For this experiment you are given a pen and a scratch paper. We kindly ask you to leave them at your place after the experiment.

Application: *(These instructions were read aloud, but not distributed)*
First, we ask you to apply yourself a nasal spray. Neither you nor the attending experimenters know whether you are given oxytocin or a non-effective aqueous solution (placebo). The sprays are distributed by a medical professional or the experimenter. Please refrain from

application until asked to do so. All participants should apply the spray at the same time. Please do not communicate with your neighbors. If you have any private questions, please raise your hand. An experimenter will join you to answer your questions.

(Instructions of each part were presented sequentially)

Part I

Questionnaire: *(These instructions were read aloud, but not distributed)*
First, we kindly ask you to fill in the questionnaire on the screen. You find the corresponding instructions also on the screen.

Part II
Interactive decision making situation

In the beginning of part II, all participants are divided into **groups of three**. Your group therefore consists of you and two other group members. At no point in the experiment you will get to know the identity of any of your group members. It is however not possible that you are grouped with the same people as in the previous part of the experiment.

Your decision in part II
In part II **you decide between two options: Option A and option B**. The other two group members also decide between the options A and B.

Your profit in part II
The profit of each group member will be calculated in the same way and is given by the following payout matrix:

	Both of your group members choose option A	**One of your group members chooses option A, the other group member chooses option B**	**Both of your group members choose option B**
You choose option A	20	32	44
You choose option B	12	24	36

This means:
If all three group members choose option B, each of them receives 36 Guilders in part II. Thus, you as well as both your group members each earn 36 Guilders in part II.
If you and a second group member would choose option B and the third member chooses option A, then you and the second group member receive 24 Guilders each, whereas the third member, who chose option A, earns 44 Guilders.
If you would choose option B and both other group members choose option A, then your profit in part II is 12 Guilders, while both other group members each earn 32 Guilders.

If all three group members choose option A, then each of them receives 20 Guilders in part II.

If you and another group member would choose option A and the third member chooses option B, then you and the second group member receive 32 Guilders, whereas the third group member, who chose option B, earns 12 Guilders.
If you would choose option A and both other group members choose option B, then your profit in part II is 44 Guilders, while both other group members each earn 24 Guilders.

Therefore, your profit depends on both your own decisions and the decisions of the other two members of your group.

In part II, you submit your decision by choosing either of the following options on the screen:

- „I choose option A“ or
- „I choose option B“

After you have decided, please click “OK” in the bottom right corner of the screen. Once you have clicked “OK”, you cannot change your decision. **Information** about the decision of the other two group members and your resulting profit in part II will be shown at the end of the experiment after part III.

Control questions

Before we start with part II, we ask you to work through the sample exercises on the screen. If you have any questions, please raise your hand. We will then come to you and answer your questions privately. You are given 45 seconds for solving each exercise. If you have not finished an exercise until the time is up, the screen will disappear without any further consequences. Therefore, please work quickly.

Part III
Second interactive decision making situation

In the beginning of part III, all participants are again divided into **groups of three**. Your group therefore consists of you and two other group members again. As before, at no point in the experiment you will get to know the identity of your group members. It is, however, not possible that you are grouped with the same people as in the previous parts of the experiment.

Part III consists of **two stages**. The first stage works like part II: Again, you **choose one of the two options A or B**. In the second stage, first you are informed about the decisions of the other two group members in the first stage. Then, in the second stage, you **decide whether you want to reduce the first stage profit of the other two group members** and, if so, by how much. This is done by distributing deduction points. In the following we describe part III in detail.

<u>First stage</u>

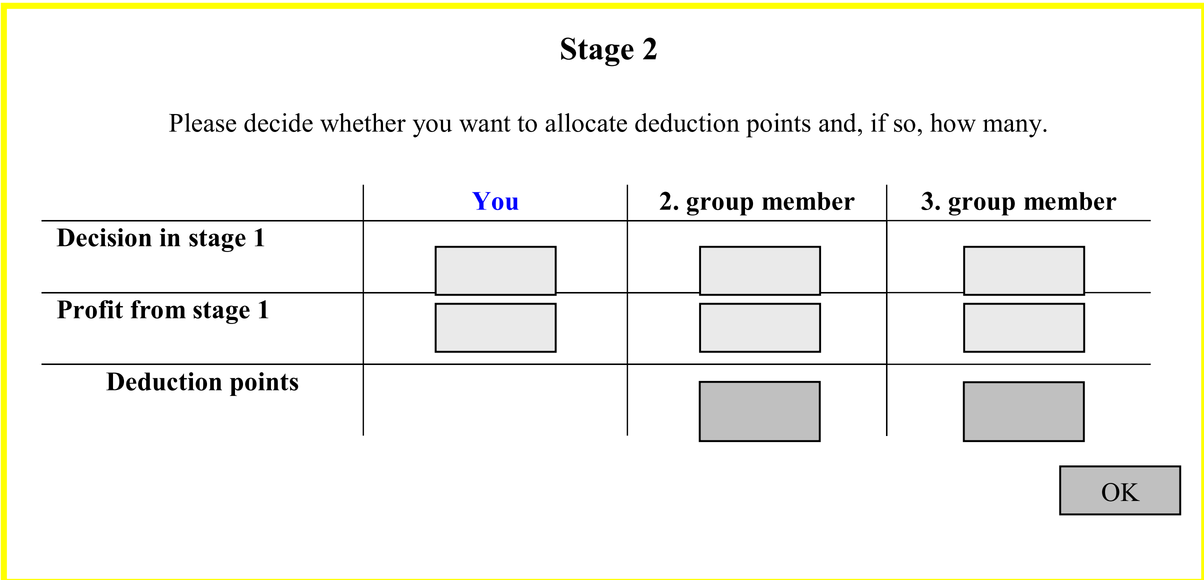
The first stage works like part II: Again, you **choose one of the two options A or B**. The payout calculation matrix is the same as in part II:

	Both of your group members choose option A	**One of your group members chooses option A, the other group member chooses option B**	**Both of your group members choose option B**
You choose option A	20	32	44
You choose option B	12	24	36

This means: If all three group members, including you, choose option B in the first stage, then all three earn 36 Guilders in the first stage and so forth. The decision screen is the same as in part II. Please click "OK" once you have made a decision.

<u>Second stage</u>

Once all members of your group have come to a decision in the first stage, you receive feedback in the second stage on the decisions of the other two group members. In this stage, you can **reduce the income of every other group member or leave it unchanged by distributing deduction points. The other group members can reduce your profit in the same manner** if they wish to. This can be seen in the following screen of the second stage:



In this screen you are shown how each group member decided in the first stage between option A and B (see first row in the screen shot: "Decision in stage 1"). Your own decision is displayed in the left column in blue, while the decisions of the other two group members are displayed in the remaining two columns. In the second row you see the resulting income for each group member ("profit from stage 1").

Now for each of the other two group members, you decide if and how many deduction points you want to distribute to each of them. To do so, in the third row (“Deduction points”), you enter an **integer between 0 and 10** in the corresponding field (0 and 10 included). If you do <u>not</u> want to reduce a group member’s income, enter 0 deduction points. Please note: **You can distribute a maximum of 10 deduction points to each group member**.

If you distribute deduction points to a group member, **you will reduce his income by triple the amount of deduction points you distributed**. For example, if you distribute one deduction point to another group member (by entering “1” in the corresponding field on the screen), you reduce his income from the first stage by 3 Guilders. If you distribute 2 deduction points to a group member, you will reduce his income by 6 Guilders and so forth. Therefore, each deduction point you distribute to a group member reduces the first stage income of that group member by 3 Guilders. Likewise, your first stage income is reduced by 3 Guilders for every deduction point distributed to you.

Cost of received deduction points (in Guilders) = 3 * sum of received deduction points

Distributing deduction points costs you Guilders. Every deduction point you distribute costs you **one Guilder**. For example, if you distribute in total 11 deduction points, this will cost you 11 Guilders and so on. Of course, if you do not distribute any deduction points you do not have to bear any costs. Please note: Once you have clicked “OK”, you cannot change your decision anymore.

Your profit at the end of the second stage, this means your profit in part III, is calculated as follows:

Your profit at the end of stage two (in Guilders) =
= Profit from first stage – 3 * (sum of received deduction points) – sum of distributed deduction points

If, for example, your profit from the first stage is 20 Guilders and you receive a total of 3 deduction points from both other group members and you distribute one deduction point yourself to one other group member, then your period income is calculated by 20 – 3*3 – 1 = 10 Guilders. However, if your income from the first stage is 12 Guilders, you receive 2 deduction points by the other group members and distribute none yourself, then your income at the end of the second stage is calculated by 12 – 2*3 – 0 = 6 Guilders.

Please note: **The profit of any group member can never be reduced further than to 0 Guilders by the other two group members**. I.e. when the cost from received deduction points for a group member are higher than his first stage income, then his profit is reduced to 0 Guilders and not further. However, even in this case the group member still has to carry the cost for the deduction points distributed by him.

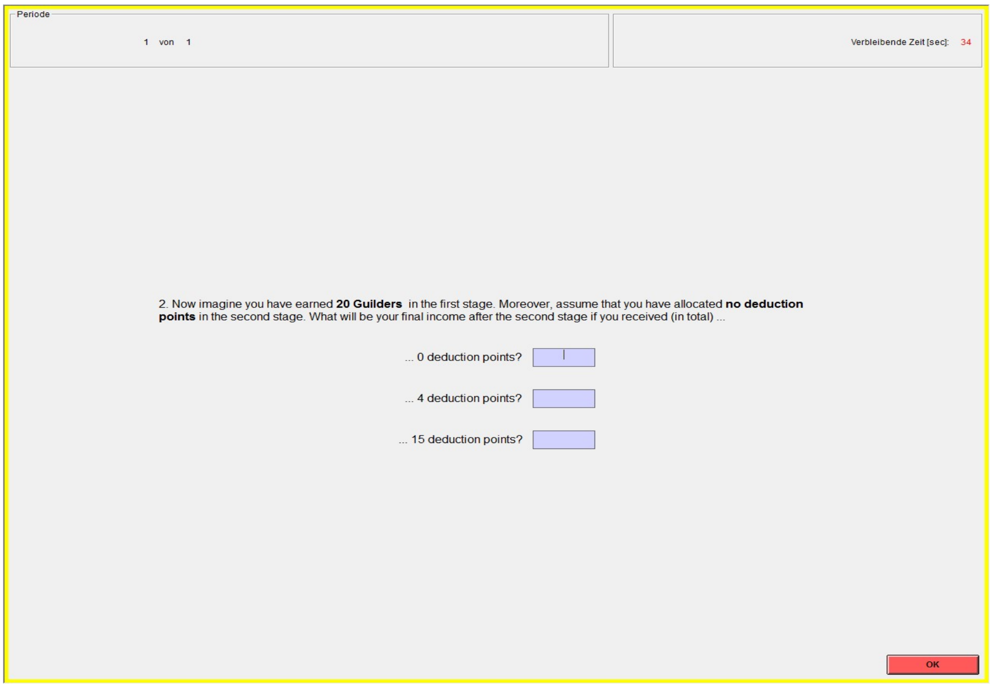
This means that your earnings in part III (in Guilders) may be negative at the end of the second stage, if your cost for distributing deduction points are higher than your (possibly reduced) income from the first stage. **However, these losses can certainly be avoided by your own decisions**. Possible losses are covered by the 20 Guilders which you receive automatically at the beginning of the experiment and which will be charged at the end against your total income from the experiment.

Information on how many deduction points the other two group members distributed to you in total and on your profit in part III will be given to you at the end of the experiment after part IIII.

Control questions

Before we begin with part III, we ask you to work through the sample exercises on the screen. If you have any questions, please raise your hand. We will come to you and answer your questions privately. You are given 45 seconds for solving each exercise. If you have not finished an exercise until the time is up, the screen will disappear without any further consequences. Therefore, please work quickly.

(Screenshot of control questions in part III)



Part IV

Questionnaire

Please imagine again the situation from part III when you were shown the decisions of the other two group members between option A and B, but before you have chosen your distribution of deduction points. For this, the following screen shows the decision each group member made in the first stage of part III:

Questionnaire

	You	2. group member	3. group member
Decision in stage 1			
Profit from stage 1			
Your emotions		1. I am angry. I disagree completely ○ ○ ○ ○ ○ ○ ○ I agree completely 2. I am grateful. I disagree completely ○ ○ ○ ○ ○ ○ ○ I agree completely 3. I feel guilty. I disagree completely ○ ○ ○ ○ ○ ○ ○ I agree completely 4. I am happy. I disagree completely ○ ○ ○ ○ ○ ○ ○ I agree completely 5. I am confused. I disagree completely ○ ○ ○ ○ ○ ○ ○ I agree completely 6. I am ashamed. I disagree completely ○ ○ ○ ○ ○ ○ ○ I agree completely 7. I am surprised. I disagree completely ○ ○ ○ ○ ○ ○ ○ I agree completely 8. I am disappointed. I disagree completely ○ ○ ○ ○ ○ ○ ○ I agree completely	1. I am angry. I disagree completely ○ ○ ○ ○ ○ ○ ○ I agree completely 2. I am grateful. I disagree completely ○ ○ ○ ○ ○ ○ ○ I agree completely 3. I feel guilty. I disagree completely ○ ○ ○ ○ ○ ○ ○ I agree completely 4. I am happy. I disagree completely ○ ○ ○ ○ ○ ○ ○ I agree completely 5. I am confused. I disagree completely ○ ○ ○ ○ ○ ○ ○ I agree completely 6. I am ashamed. I disagree completely ○ ○ ○ ○ ○ ○ ○ I agree completely 7. I am surprised. I disagree completely ○ ○ ○ ○ ○ ○ ○ I agree completely 8. I am disappointed. I disagree completely ○ ○ ○ ○ ○ ○ ○ I agree completely
Your fairness judgment		The second member's decision was very unfair ○ ○ ○ ○ ○ ○ ○ very fair	The third member's decision was very unfair ○ ○ ○ ○ ○ ○ ○ very fair

OK

In the first row ("Decision in stage 1") the decisions in your group from the first stage of part III are shown again. The second row ("Profit from stage 1") shows the corresponding profits from the first stage. The third row contains **eight statements** for each group member which could describe your emotion states towards the second and third group member. For each statement, please indicate separately how you felt towards group member two and three, respectively, and **before you had decided on the allocation of your deduction points**. For this, you are provided with a seven-point scale ranging from “I disagree completely” to “I agree completely”.

In the last row (“Your fairness-judgment”), please indicate how you rate the fairness of the decisions of each group member, in the first stage (see first row). To do so, you are provided with a seven-point scale ranging from “very unfair” to “very fair”. When you are finished with your assessment, please click “OK” in the bottom right corner of the screen. For the completion of the questionnaire you will receive **10 Guilders**.

Debriefing and payout

Once all participants have completed the last questionnaire, the outcome of all individual and interactive decisions will be revealed. Additionally you will learn your profits from all parts of the experiment as well as your total profit. After that, you will receive your payment from the experiment individually and in cash.

Chapter 2

Oxytocin promotes dishonesty under competition*

2.1 Introduction

There is a notion that keen competition on markets may lead to an erosion of moral values (Falk & Szech, 2013; Faravelli et al., 2015; Shleifer, 2004). Such a potential incentive to engage in unethical behavior questions the positive welfare view on perfect markets. The underlying idea is that competition could force companies to engage in the same unethical behavior as their competitors, because else, they would be driven out of the market (Shleifer, 2004). In a famous example, a large German company paid a $ 1.6 billion fine for bribing government officials in numerous countries (Schubert & Miller, 2008, December 20), and one of the accused managers argued that bribery was usual in the business the company was operating in. If this was true, not only an intrinsic motivation (such as greed), but also a competitive environment could promote immoral (and illegal) behavior. Actually, the reference to competitors is often used as a justification for own unethical behavior. If the belief on competition and (un)ethical behavior is shared among market participants, unethical behavior can become a 'norm' in its own and may thus be considered as little harmful or even appropriate (Falk & Szech, 2013; Gino et al., 2009). Recent empirical evidence confirms the notion that the mere presence of competition may render unethical behavior as an acceptable measure (Belot & Schröder, 2013; Falk & Szech, 2013; Faravelli et al., 2015; Schwieren & Weichselbaumer, 2010).

* This chapter is based on joint work with Andrea Jobst, Martin Kocher, Fabian Loy and Norbert Müller.

Surprisingly, there is still scarce empirical evidence regarding the sources of accepting unethical behavior like dishonesty in competitive environments. Here, we argue that competition using unethical measures constitutes an environment that contains an element of in-group conformity, as all participants accept this situation and all participants expect others to use the same unethical measures as well. In other words, an alternative social norm is created and a stronger preference for in-group conformity could reinforce unethical behavior, once the belief on engaging in unethical behavior is shared sufficiently widely. Not surprisingly, a frequently used justification in competitive environments can be summarized as 'everyone does it'. This would also explain the persistence and prevalence of unethical behavior in sports (such as doping), business (such as fraudulent accounting) and politics (such as using lies in electoral competition), despite the existence of regulations, rules and harsh punishments against misconduct (Faravelli et al., 2015).

The roots of in-group conformity of social beings, like humans, lies in the evolutionary advantage of sharing common behavior, opinions and knowledge within a community, which ultimately increases survival likelihood of each individual within the group (Cialdini & Goldstein, 2004). We argue that the evolved neurobiological circuits, which sustain and motivate in-group conformity, may also be involved in promoting unethical conduct in competitive environments when unethical behavior becomes the 'new' norm. If this is true, an exogenous variation of the preference for conformity would lead to a change in the prevalence of unethical conduct in the presence of competition.

In order to achieve this variation, we link unethical behavior in competition to oxytocin, an endogenous neuropeptide produced in the mammalian hypothalamus, that has been shown to play an important role in group affiliation (Donaldson & Young, 2008b) and in-group conformity (Stallen et al., 2012). In particular, Stallen et al. (2012) provided evidence that subjects given oxytocin showed a higher inclination to conform their judgments to others when they were asked to rate the attractiveness of novel visual stimuli. However, this in-group bias was absent in subjects treated with placebo. Moreover, there is evidence that oxytocin has a significant discriminatory effect in moral judgments, as subjects favored their in-group over an out-group in the commonly used trolley paradigm (De Dreu et al., 2011). In a similar vein, subjects under the influence of oxytocin showed a higher propensity to cooperate with in-groups than with out-groups (De Dreu et al., 2010) and would even lie for the benefit of their in-group (Shalvi & De Dreu, 2014).

Taken together, these findings suggest that oxytocin is involved in the modulation of human's preferences for conformity. Due to its positive effects on in-group conformity, we hypothesized that oxytocin leads to an increase of unethical conduct in a competitive environment when it is believed that one's peers engage in the same unethical behavior. If this belief is not present in a non-competitive environment, oxytocin should not have any effect on unethical behavior. In particular, we want to test whether an exogenous variation of brain oxytocin levels modulates the inclination to break (or bend) a widely accepted norm, i.e. honesty, in a competitive task compared to a non-competitive situation. In the interest of parsimony, we exclude the possibility that unethical conduct can be monitored or detected on the individual level.

2.2 Experimental design

We tested our hypothesis in a double-blind and placebo-controlled experiment, in which 120 healthy male subjects were randomly assigned to either the oxytocin group (N = 60) or to the placebo (N = 60) group. Subjects self-administered a nasal spray which contained a single intranasal dose of 24 IU of oxytocin or a placebo. Then, 60 minutes after intake, subjects received a coin and performed two different versions of the coin tossing task that has been shown to reliably measure dishonesty (Abeler, Becker, & Falk, 2014; Cohn, Fehr, & Marechal, 2014; Houser, Vetter, & Winter, 2012; Shalvi & De Dreu, 2014). Subjects were told to privately flip the coin, and were paid according to the reported results. The advantage of this task lies in the fact that subjects do not face the risk of detection as we were credibly not able to observe lying of a particular subject. However, by comparing the reported outcomes of all individuals with their statistical chance implied by a fair coin, we were able to assess honesty on an aggregate level.

In part one of our experiment (the non-competitive task), subjects received a € 1 coin and were told to flip the coin three times, but to report the outcome of the first two tosses.[1] For each reported 'tail' they would win a real monetary prize of € 1.66 (for details see *Methods*). Part one serves as a baseline, as it measures participants' general inclination to break a norm (i.e. lying) for a financial benefit without directly considering others' behavior. That is, subjects have no peer to compare with, as others' behavior is not rendered salient in the non-competitive task. Therefore, the non-competitive task constitutes an ideal tool to analyze behavior when no enhanced element of conformity is present.

In part two (the competitive task), subjects were also asked to flip a coin three times, and again, to report only the first two outcomes. To induce a competitive environment,

[1] This method is commonly used, as it gives subjects the opportunity to lie by reporting desired counterfactuals (i.e. not the relevant but the best outcome of all tosses) and treat them "as if" they really happened.

participants' payment in this part depended on their report and on the report of a randomly matched subject in their session. In particular, the subject who reported the highest number of tosses with 'tails' won € 3.33. In case of a tie, the prize was split equally. This part involves a strategic element, as subjects may want to form beliefs about the actions of others if they care about the norm at all. Thus, immediately after both parts we elicited subjects' beliefs regarding the frequency of reported 'tails' in their experimental session.

Compared to the first part, in part two actions of others are rendered salient as subjects were (randomly) matched to each other and had to interact in a strategic setting. By comparing the effect of oxytocin on behavior in the strategic (competitive) with the non-strategic (non-competitive) task, we can disentangle the effect of oxytocin on general dishonesty and the oxytocin-induced effect of conformity on lying behavior.

As empirical evidence suggests that lying causes psychological costs due to the loss of a positive self-image (Abeler et al., 2014; Cohn et al., 2014; Fischbacher & Föllmi-Heusi, 2013; Mazar, Amir, & Ariely, 2008; Shalvi, Eldar, & Bereby-Meyer, 2012), we additionally asked participants to rate eight general emotions on a seven-point Likert scale at the end of the experiment (Hopfensitz & Reuben, 2009).[2] This should enable us to indirectly detect a bad conscience associated with dishonest reports. Due to a decrease in psychological lying costs in the competitive environment, we conjectured that subjects exhibit more negative emotions related to over-reporting in the non-competitive task compared to the competitive task.

2.3 Results

2.3.1 Analysis of the reported outcomes in the coin tossing task

As the flip of the coin was private, we measured honesty on an aggregate level by comparing reported tosses with their expected frequency based on a fair toss. As can be seen in Fig. 2.1, participants' behavior strongly deviated from honest reporting, regardless of condition or treatment. Overall, participants over-reported successful tosses (i.e. tossing 'tails' twice) compared to its statistical chance in the non-competitive condition (65%; binomial $z = 10.11$; $P < 0.001$) and in the competitive condition (80.8%; binomial $z = 14.12$; $P < 0.001$). A comparison of both conditions reveals that dishonesty was more pronounced in a competitive compared to a non-competitive environment (Wilcoxon matched-pairs test; $z = 2.237$; $P = 0.025$). In short, subjects tended to lie more when competition was present.

[2] We did not elicit emotions immediately after each task to avoid a spillover effect of the questionnaire on the subsequent tasks, as we asked for feelings such as shame and disappointment.

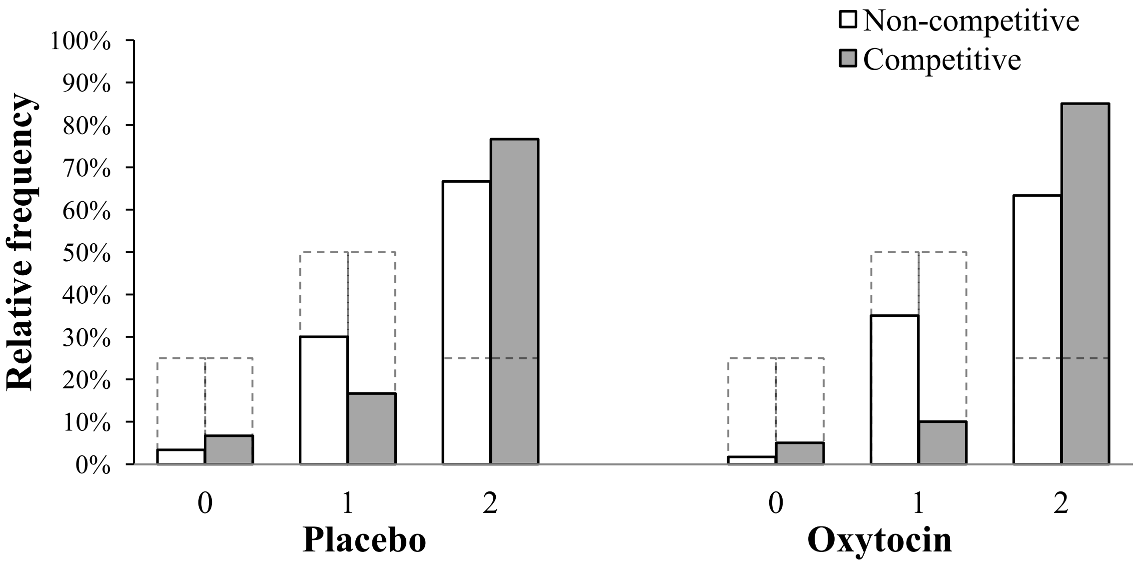


Fig. 2.1 | Distribution of reported 'tails' as a function of treatment (oxytocin vs. placebo) and condition (competitive vs. non-competitive). The dashed bars depict the outcome by a fair toss (i.e. honest reporting).

However, this effect was almost completely driven by subjects given oxytocin. That is, participants treated with placebo barely reacted to a competitive environment (Wilcoxon matched-pairs test; z = 0.724; P = 0.469). Conversely, in the oxytocin group we found a substantial change to, on average, more dishonest reports in the competitive environment, as the difference between both conditions is significant (Wilcoxon matched-pairs test; z = 2.315; P = 0.021). As a result, subjects under the influence of oxytocin showed a significantly higher inclination to break a norm (here: honesty) under competition compared to a non-competitive setting.

2.3.2 Computed proportions of lying types

Previous empirical findings indicate that there may be two types of dishonesty: some subjects lie to the full possible extent in order to maximize profits, whereas some subjects only partially lie to preserve a certain level of positive self-image (Fischbacher & Föllmi-Heusi, 2013; Gino et al., 2009; Houser et al., 2012; Utikal & Fischbacher, 2013). In our experiment subjects had the opportunity to report honestly, lie partially, or lie fully about the outcome of their coin tosses. Hence, we defined α_1 as the proportion of partial liars and α_2 as the proportion of full liars in our sample. As full liars would always report '2', their probability of reporting this outcome is one. In contrast to full liars, participants who lied partially are assumed to improve their outcome by exactly one Hamming distance, and reported '1' instead of '0' or '2' instead of '1'. By comparing the reported frequencies $f(x)$ of a particular outcome $x \in \{0, 1, 2\}$ with their ex ante statistical probabilities, we were able to compute an estimate for the actual frequency of each type as a function of the observed frequencies of each outcome (see appendix B1 for proof):[3]

Proportion of partial liars: $\alpha_1 = 4 \cdot (f(1) - 2 \cdot f(0))$

Proportion of full liars: $\alpha_2 = \frac{4}{3}(f(2) - 2 \cdot f(1) + 4 \cdot f(0) - 0.25)$

Table 2.1 shows the computed proportions of each type in the competitive environment. Accordingly, a comparison between treatment groups in the competitive task reveals that more lying was present in the oxytocin group (Fisher's exact-test; $P = 0.003$), such that the neuropeptide increased full lying to 80%, compared to 60% in the placebo group. Interestingly, partial lying completely vanished in the oxytocin group when comparing the competitive with the non-competitive task. This indicates that subjects, who partially lied to benefit from dishonesty but still wanted to preserve a positive self-image, became completely dishonest under the influence of oxytocin. Thus, it seems that an intake of oxytocin results in a drop in psychological costs of lying, supposedly making lying more ethically permissible under competition. Combined with the finding that oxytocin showed no effect on lying in the non-strategic task, we conclude that the effect of oxytocin on dishonesty was limited to the competitive setting.

	Proportions of lying types	
	Placebo group	*Oxytocin group*
Truth-tellers	0.26	0.20
Partial Liars (α_1)	0.13	0
Full Liars (α_2)	0.60	0.80

Table 2.1 | Proportions of lying types in the oxytocin and in the placebo groups under the competitive environment

2.3.3 Reported beliefs in the coin tossing task

However, the question arises whether the adverse effect of oxytocin on honesty was rooted in a preference for conforming actions to one's peers. Such reasoning would require that

[3] Houser, Vetter, and Winter (2012) used a similar method; however, their analysis consisted of two types, liars and truth-tellers. Our work extends their notion by allowing for partial lying.

subjects anticipated more lying of their peers in the oxytocin group compared to the placebo group. To test whether our observed treatment effect is in line with subjects' beliefs about the actions of others, we elicited their beliefs regarding the lying rate in their sessions. More precisely, we asked them to guess the frequency of subjects who reported the profit-maximizing outcome (i.e. tossing 'tails' twice). In case the observed treatment effect of oxytocin was rooted in an increase of in-group conformity, we would expect more lying when subjects believed others would lie as well. Fig. 2.2 shows the average anticipated over-reporting for each condition. Remarkably, the anticipated lying rate in the competitive environment was significantly higher in the oxytocin group than in the placebo group (87.7% vs. 76.5%; Mann-Whitney U-test; $z = -2.708$; $P = 0.006$). However, there was no statistical difference between oxytocin and placebo in the non-competitive environment (83.7% vs. 81.1%; Mann-Whitney U-test; $z = -1.035$; $P = 0.300$). Thus, it seems that the change in (dis)honesty in the oxytocin group is mainly driven by the beliefs on the honesty of others. As oxytocin negatively influenced those beliefs, we argue that subjects found it ethically more permissible to lie, as long as others would lie as well.

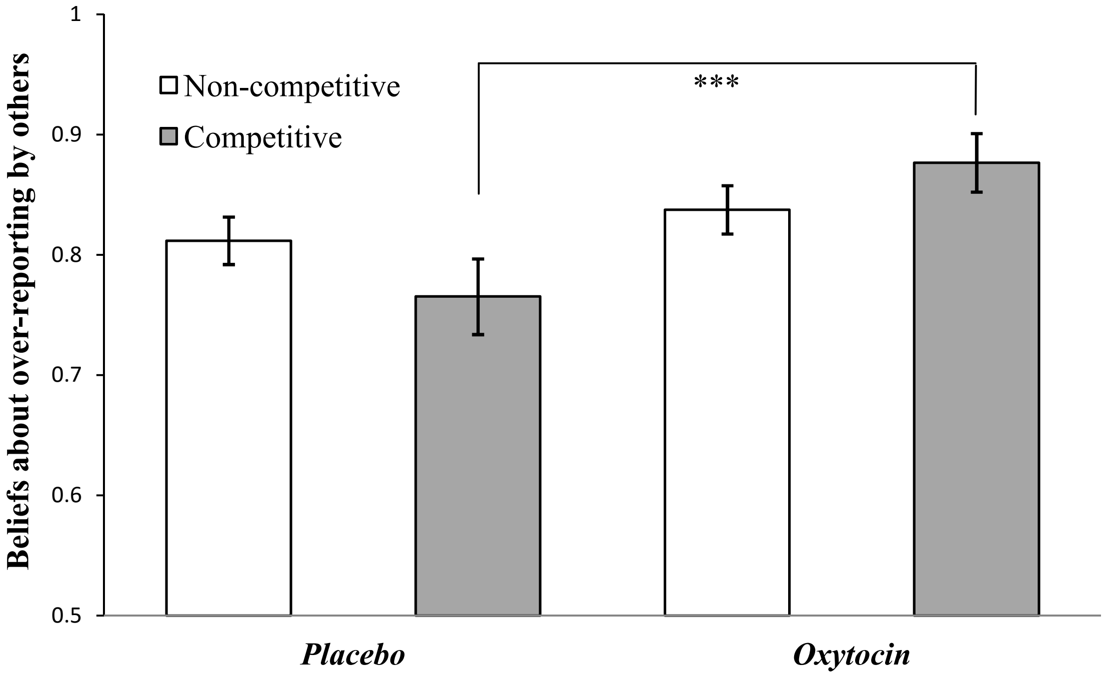


Fig. 2.2 | Beliefs about reporting 'tails' twice as a function of treatment (oxytocin vs. placebo) and condition (Non-competitive vs. Competitive). Note that ***, **, * denote significance at the 1%, 5%, 10% level.

2.3.4 Analysis of conformity in the coin tossing task

To provide evidence that oxytocin induced in-group conformity, we have to test whether subjects aligned their actions with the actions of their peers. As subjects received no information about their peers' choices, we assume that a conformist subject would exactly do what he or she believed others would do. To elicit beliefs precisely, we asked participants to guess the reported frequency $g(x)$ of each outcome x within a session. Accordingly, we were able to compute the believed average reported outcome. That is, we calculated for each subject his or her beliefs about the mean of all reported tosses b_{ij} within a session in the following way:[4]

$$b_{ij} = \sum g(x) \cdot x.$$

Furthermore, in-group conformity requires the coherence of a subject's action with his/her beliefs about his/her peers' actions. This suggests that the difference between beliefs about cheating rates and actual cheating behavior should be smaller in the oxytocin group than in the placebo group if oxytocin caused in-group conformity. To test this, we computed an inconformity index λ_i for each subject i, and calculated the absolute difference of an actual decision a_i of subject i and his or her belief about the mean of all reported outcomes b_{ij} within the respective session j. Hence, the inconformity index is defined as follows:

$$\lambda_i = |a_i - b_{ij}|$$

Thus, the inconformity index λ_i increases proportionally with the divergence of one's actions from the believed choices of one's peers, and a value of 0 indicates a minimum of inconformity. Or put it differently, λ_i has a value of 0 if a subject chose what he or she believed others would choose. Our empirical analysis of inconformity levels λ_i reveals that subjects in the oxytocin group exhibited a significantly lower difference between actions and beliefs than in the placebo group in the competitive environment (oxytocin: *mean* = 0.19 vs. placebo: *mean* = 0.26; Mann-Whitney U-test; $z = 2.120$; $P = 0.034$). The result is in line with the notion that oxytocin induced conformity to the (believed) actions of peers. However, in the non-competitive situation when strategic considerations played no role, both groups did not differ significantly in terms of the inconformity index (oxytocin: *mean* = 0.36, placebo: *mean* = 0.31; Mann-Whitney U-test; $z = -0.621$; $P = 0.534$). That is, in non-strategic situations subjects under the influence of oxytocin showed no different levels of inconformity compared to placebo. The absence of oxytocin-induced conformity in the non-competitive task is not surprising; we did not render subjects' peers salient in the non-competitive task. However, this changed in the competitive environment, since actions of others became much

[4] By definition, b_{ij} is restricted to the interval [0; 2]. Thus, beliefs are restricted to the same interval as the actual choices in the tasks.

more salient, as subjects had to interact with each other in a strategic setting. We argue that this inherent difference between the two conditions has diverted a subject's attention from herself to her peers.

2.3.5 Analysis of emotions related to over-reporting

The question arises why the oxytocin-induced conformity had a strong effect on the inclination to lie. One possible explanation could be a change in psychological costs, as subjects perceive an immoral act as more permissible when others choose to behave unethically as well. If this is true, then we should observe a significant correlation of negative emotions to subsequent dishonesty in the non-competitive task, but not in the competitive task. Table 2.2 shows that in the non-competitive environment reported tosses and beliefs about the lying rate in the respective session were good predictors for the emotion shame.[5] However, we found no such correlation in the competitive environment. Moreover, shame increased with dishonesty and decreased with one's belief about the average dishonesty level of others. Remarkably, this relationship was only present in the non-competitive environment. As shame was only correlated with behavior and beliefs in a non-strategic situation, it seems that the presence of a competitive environment leads to a crowding out of negative emotions. That is, as competition is perceived as a legitimate environment for lying, subjects had no negative emotions associated with their dishonesty. As expected, oxytocin per se had no direct effect on shame, as it should not modulate negative emotions, independently of actions and beliefs. Accordingly, we were able to rule out any general effects of oxytocin on dishonesty or on psychological costs related to unethical behavior.[6]

[5] Similar to Hopfensitz and Reuben (2009), we elicited seven other general emotions to avoid an experimenter demand effect (i.e. anger, gratefulness, guilt, gladness, irritation, surprise, disappointment) before and after the tasks. Except for the emotion 'shame', none of them showed any significant correlation with treatment or choices in the tasks.

[6] We used a regression analysis to examine correlations between over-reporting and emotions, because a comparison of emotions for a specific lying rate is not possible since our design does not allow detecting lies on the individual level.

	Dependent Variable: Δ shame	
	(1) Competitive environment	(2) Non-competitive environment
Treatment (Oxytocin=1 or Placebo=0)	0.193	0.233
	(0.213)	(0.198)
Reported outcome of tosses	0.169	0.577***
	(0.259)	(0.217)
Believed lying rate in session j	−0.442	−1.466***
	(0.539)	(0.488)
Constant	0.705	1.808**
	(0.712)	(0.731)
N	120	120
R^2	0.0104	0.0879

Table 2.2 | OLS regression of changes in the emotion 'shame'. The self-reported level of shame is used as a dependent variable in an OLS regression. Emotion ratings after the decision in the tasks were normalized by subtracting from reported baseline levels that were elicited immediately before the experiment. To avoid an experimenter demand effect we elicited seven other general emotions. Note that ***, **, * denote significance at the 1%, 5%, 10% level and standard errors are in parentheses.[7]

Since we found no associations of negative emotions with dishonesty in competition, we conclude that subjects in the oxytocin group were more conforming to their peers and faced lower psychological costs of unethical behavior. That is, the consequence of conforming to others' actions leads to a diffusion of responsibility.

2.4 Discussion

The present work examines the effect of oxytocin on a widely accepted norm, i.e. honesty, by comparing lying rates between a competitive and a non-competitive environment. In particular, we employed a coin tossing task in which the experimenter was not able to detect

[7] A full model with all independent variables in one regression exhibits exactly the same pattern (see appendix B2 for details).

lies on an individual level, but to infer dishonesty on an aggregate level in an unobtrusive way. Our results provide evidence that oxytocin induced conformity, which had a detrimental effect on honesty in the competitive environment, as lying rates significantly increased compared to a non-competitive environment. This effect was not present in the placebo group. Our results are in line with previous empirical findings that demonstrated the detrimental effects of competition on ethical behavior (Belot & Schröder, 2013; Falk & Szech, 2013; Faravelli et al., 2015; Schwieren & Weichselbaumer, 2010). Our work extends previous findings by providing evidence that in-group conformity plays a crucial role in competitive environments, as unethical behavior is perceived as more permissible under competition. Furthermore, our analysis indicates a substantial shift of partial lying to full lying due to the oxytocin treatment. Obviously, subjects with chemically-induced in-group conformity found it ethically more permissible to lie when a justification (i.e. competition and beliefs about others' behavior) was present. We conclude that subjects who partially lied to preserve a positive self-image were more prone to justifications. Moreover, we showed that competition reduced the psychological costs of lying, as subjects perceived unethical behavior as more permissible or even appropriate when they believed that their peers would do the same.

Our study is most closely related to the work of Shalvi and De Dreu (2014), in which participants self-administered oxytocin or a placebo and had the opportunity to lie for the financial benefit of their respective group. They found that an exogenous application of oxytocin increased group-serving lying in a cooperative environment, in which dishonesty generated a joint profit among group-mates. However, no effect was found when subjects lied for their individual profit. The authors argue that lying in a cooperative environment favors and supports one's in-group and serves directly the interests of their group. This result is in line with the well-documented pro-social effects of oxytocin (De Dreu et al., 2010; Kosfeld et al., 2005; Zak et al., 2007), as mutual beneficial dishonesty may be rooted in reciprocal motives.[8] That is, confirming to group-serving dishonesty may be more acceptable, as it is mutual beneficial. However, in our study we still find an effect of oxytocin-induced in-group conformity in a competitive environment, although by design lying generates negative externalities to one's peers. Thus, in our setting (positive) reciprocal motives cannot play a role. Since the detrimental effect of oxytocin on honesty persists in a competitive as well in a cooperative situation, we argue that conformity to others' actions might be the main driver of the observed effect of oxytocin on dishonesty.

[8] However, the question is still not answered why oxytocin induced reciprocal motives in their gains treatment but not in their loss treatment, as McDermott, Fowler, and Smirnov (2008) point out that avoiding losses seem more fundamental for the survival of a group.

Furthermore, our results demonstrate that competition may serve as a form of justification for unethical behavior. In a similar vein, Shalvi et al. (2012) showed that a lack of justification may increase honesty in a simple lying task. They argue that justifications enable people to lie and preserve a positive self-image at the same time. A similar logic applies to markets with negative externalities on third parties. Falk and Szech (2013) argue that market participants may perceive themselves as not pivotal and might therefore trade unethically produced goods (e.g. child labor), despite their general objection to it. Again, the justification of being negligible in a (large) market drives immoral behavior. Furthermore, Gino et al. (2009) found that the inclination to lie was significantly higher when subjects observed dishonest behavior of an in-group peer, and lower when subjects observed the same behavior of an out-group member. The authors argue that, according to social identity theory, subjects distanced themselves from unethical out-group behavior to maintain a positive social identity and self-image, and at the same time perceived the identical immoral behavior as more permissible if it was socially accepted by their peers. Similarly, a negative business culture may have a detrimental effect on honesty of employees who would be honest otherwise (Cohn et al., 2014). Consequently, it is straightforward to presume that a justification may suffice to reduce psychological costs of unethical behavior. Here, we argue that in our experiment competition served as a justification and therefore eroded moral values, as subjects conformed to the unethical conduct of their peers.

As competition may decrease psychological costs of dishonesty, we would expect that dishonesty in a competitive environment would not cause negative emotions compared to a non-competitive situation. That is, subjects in the competitive environment had the possibility to justify their actions by arguing ‘everyone does it’ and to share their own responsibility with their peers. As a result, subjects in the competitive environment should be less prone to negative emotions caused by a bad conscience. This notion is in line with our results. Indeed, we observed that lying in a non-competitive environment was a good predictor of subsequent negative emotions, in particular shame, whereas this relationship completely vanished in the competitive task. We conclude that subjects were able to avoid a bad conscience by diffusing the responsibility to their peers, which was not possible in the individual task. Accordingly, it seems to be easier to preserve one’s positive self image, if one can attribute her behavior to the circumstances and to one’s peers.

Despite the provided evidence regarding the positive effects of oxytocin on conformity, one could also come up with the notion that the observed treatment effect of oxytocin could be rooted in a better strategic assessment of one’s peers. And indeed, we observed that subjects given oxytocin were better in predicting the actions of others (see supplementary analysis in appendix B3). Therefore, it could be the case that oxytocin enhanced the ability to think strategically. The more accurate assessment of one's peers and their likelihood to lie could have caused the change in behavior. In other words, oxytocin might have promoted strategic thinking, which is why subjects were closer to the selfish Nash equilibrium of the game. However, we argue that if this was true, subjects should still bear the same psychological costs caused by behaving dishonestly. However, our data provides evidence that subjects exhibited no negative feelings related to lying in the competitive environment, and thus showed no signs of a bad conscience or remorse. We argue that this lack of remorse can only be explained if subjects exhibited lower psychological costs of lying.

2.5 Conclusion

Our results suggest that it is not only necessary to prevent unethical conduct per se, but to prevent the notion that unethical behavior is standard practice. That is, the mere belief that others would use unethical measures has often a detrimental effect on honesty due to in-group conformity. In extreme cases, tax avoidance would be perceived as widespread, corruption would be seen as a common part in business activities and performance-enhancing drugs as a normal component of one's training. In light of this, we argue that policies and regulations should tackle unethical behavior not only by employing deterrence measures like punishment. We rather want to emphasize that appealing to one's self-image could increase psychological costs related to unethical conduct, as several empirical findings demonstrated that most people care about their positive self-image (Abeler et al., 2014; Cohn et al., 2014; Fischbacher & Föllmi-Heusi, 2013; Mazar et al., 2008; Shalvi et al., 2012).[9] Since there is evidence that even a simple message of approval (or disapproval) of others has a significantly positive effect on energy savings of households (Schultz, Nolan, Cialdini, Goldstein, & Griskevicius, 2007), we argue that appealing to one's conscience might have a similar impact on ethical behavior as well (Mazar et al., 2008). Unethical behavior in negative business environments could potentially be tackled by appealing to one's personal responsibility (Mazar & Aggarwal, 2011), by promoting a positive professional identity (Cohn et al., 2014), and by rendering others disapproval to unethical behavior more salient (Schultz et al., 2007). Consequently, an open and transparent business culture with an emphasis on professional ethics is certainly able to overcome loopholes of convenient justifications, when monitoring is not possible or too

[9] Utikal and Fischbacher (2013) provided empirical evidence that subjects who wanted to preserve a positive self-image even lied to their own disadvantage.

costly. Our work contributes to previous findings by showing that it is crucial to avoid a climate of 'everyone does it'.

2.6 Methods

2.6.1 Participants

We recruited a total of 120 male subjects from different universities in Munich via Orsee (Greiner, 2004). Participants were between 18 and 35 years old (mean age ± s.d., 23.125 ± 3.32 years) and took part in seven experimental sessions at the Munich Experimental Laboratory (MELESSA). Exclusion criteria for participation were significant medical or psychiatric disorder, medication, smoking more than 15 cigarettes per day, and drug or alcohol abuse. Subjects were instructed to abstain from alcohol, smoking and caffeine for 24 h, and from food for 1 h before the experiment. At the time of recruitment participants were informed that the experiment is about the influence of oxytocin on economic decision making. The study was approved by the ethics committee of the medical department at the University of Munich. All participants provided written, informed consent before participation.

2.6.2 Substance administration

We used a double-blind, placebo-controlled study design to compare behavior in a group of subjects that received a single intranasal dose of 24 IU oxytocin (Syntocinon-Spray, Sigma-Tau; 3 puffs per nostril, each with 4 IU of oxytocin) with that of subjects in a control group that received a placebo. Subjects were randomly assigned to both groups and received either oxytocin or a placebo about 60 minutes before the relevant part of the experiment started. In one experimental session either oxytocin or a placebo was administered. The placebo contained all inactive ingredients except for the neuropeptide (Born et al., 2002; Gossen et al., 2012).

2.6.3 Behavioral experiment and questionnaires

We employed a fully computerized experimental design, whereas all decisions were made privately and anonymously. First, subjects received neutrally written instructions, which we read aloud (see appendix B4 for details). They were informed that the experiment consisted of several parts which were independent of each other. Each part had to be accomplished by all participants in order to proceed to the next part of the experiment. 60 minutes after substance administration we started with the incentivized experiment, which consisted of a non-competitive and a competitive coin tossing task. In both coin tossing tasks, subjects were asked to privately toss a coin. To ensure that it was common knowledge that the detection of lies was not possible, we instructed the participants to make sure that neither the experimenter nor any other participant could see the outcome of their coin flip. Thus, the experimenter could only infer the lying rate on an aggregate level (i.e. of all subjects), but not on an individual level. We employed a 2 x 2 study design, i.e. we used a between-subject design to analyze the effect of an exogenous variation of oxytocin on behavior in both tasks and a within-subject design to compare changes in behavior between tasks. In the non-strategic coin tossing task, we asked subjects to privately flip a coin and report the outcome of the first two tosses out of three trials. They were told that only the first two tosses are payoff relevant and for each reported 'tails' they would earn 5 MU. This method is commonly used and gives subjects the opportunity to lie only a bit and maintain a positive self image at the same time, as it has been shown that subjects report desired counterfactuals (i.e. not the relevant but the best outcome of all tosses) and treat them "as if" they happened. This generates an acceptable loophole for maintaining a positive self image (Fischbacher & Föllmi-Heusi, 2013; Shalvi et al., 2012). Then, participants' beliefs about the lying rate in their respective session were elicited by asking them to indicate the distribution of reported 'tails'. For the correct prediction subjects earned 5 MU in addition.

Then, subjects were asked to perform the competitive coin tossing task. Again, they were asked to flip the coin three times and only report the outcome of the first two tosses. However, this time they would earn 10 MU only if they would report 'tails' more often than another randomly matched participant in the respective session. As this task was a zero-sum game, we can ignore any efficiency concerns of subjects. Thus, the only difference to the first task is that a subject's payoff depends on one's own decision and on the decision of another randomly matched participant. We exclude monitoring (or punishment), to focus on psychological costs of dishonesty. Consequently, differences in dishonesty levels can therefore be traced back to an alteration of psychological costs. Then, we elicited beliefs regarding frequencies of reported 'tails' with an incentivized question.

The order of both tasks was not randomized, as we did not want to draw the attention of subjects on the cheating behavior of others.[10] That is, if we would start with the competitive coin tossing task, cheating behavior of others would be more salient as participants would have to consciously form beliefs about the (dis)honesty of others. This would in turn bias the behavior in the subsequent individual coin tossing task, as lying is more accepted and prevalent when it is believed that others do it as well (Gino et al., 2009). As a consequence, in

[10] An alternative would be to use a full between-subjects study design and to employ independent observations in all four conditions. However, we argue that a larger sample size (of at least 240) with the respective additional expenses would not justify the added value of such design.

50% of our observations we would have faced a spill-over effect and would most certainly have had an increased dishonesty level in the individual coin tossing task due to the order. Thus, there was a tradeoff between a randomized order, which would certainly lead to a substantial spillover effect on individual lying rates, and a fixed order, which would ignore possible learning effects due to our treatment. However, to the best of our knowledge, no study found a significantly positive effect of oxytocin on learning in games (Bartz et al., 2011; Baumgartner et al., 2008). Therefore, we decided to use a fixed order to exclude spill-over effects that would be caused by the salience of others' dishonesty.

To measure the emotional states of subjects throughout the experiment and especially after the coin tossing tasks, we additionally asked subjects to rate eight emotions on a 7-point Likert scale. Subjects had to assess positive and negative emotions, namely anger, gratefulness, guilt, joy, irritation, shame, surprise, and disappointment (Hopfensitz & Reuben, 2009). The rating was conducted twice, immediately after substance administration and immediately after the coin tossing tasks. The whole experiment took about 90 minutes and was programmed using z-Tree software (Fischbacher, 2007). Each MU earned in the experiment was worth 0.33 €. The final payoff consisted of the sum of incomes from the different parts and a flat fee of 4 € for participation.

Appendix B

B1. Computation of partial lying and absolute lying rates

Subjects reported the outcome of two tosses in each condition (i.e. the competitive and non-competitive environment). Thus, subjects had the possibility to report honestly, lie partially, or lie fully about the outcome of their tosses. That is, in contrast to full liars, participants who lied partially just improved their outcome by exactly one Hamming distance, and reported '1' instead of '0' or '2' instead of '1'.

By comparing the reported frequencies $f(x)$ of a particular outcome $x \in \{0, 1, 2\}$ with their statistical probabilities, we were able to compute an estimate of each type. Let α_1 denote the proportion of all partial liars and α_2 the proportion of all full liars, then the frequency of each outcome can be described by the following three equations:

I: $f(2) = \alpha_2 \cdot 1 + (1 - \alpha_1 - \alpha_2) \cdot 0.25 + \alpha_1 \cdot 0.75$

II: $f(1) = (1 - \alpha_1 - \alpha_2) \cdot 0.5 + \alpha_1 \cdot 0.25$

III: $f(0) = (1 - \alpha_1 - \alpha_2) \cdot 0.25$

As full liars always report '2' their probability of reporting this outcome is one. However, a partial liar will shift her outcome by exactly one Hamming distance. As their statistical chance of obtaining '0' is 0.25, '1' is 0.5 and '2' is 0.25, then shifting by exactly one Hamming distance yields a chance for reporting '1' of exactly 0.25, and for '2' of 0.75. Honest players will report the true outcome and will exhibit therefore the statistical frequencies of a fair coin. Solving these three equations yields the expected frequencies of each type as a function of the reported tosses.

Proportion of partial liars: $\alpha_1 = 4 \cdot (f(1) - 2 \cdot f(0))$

Proportion of full liars: $\alpha_2 = \frac{4}{3}(f(2) - 2 \cdot f(1) + 4 \cdot f(0) - 0.25)$

Proof:

By multiplying equation III by 2 and plugging it in to equation II, we obtain:

$$f(1) = 2f(0) + \alpha_1 \cdot 0.25$$

Then, by solving for α_1, we obtain the proportion of partial liars:

$$\alpha_1 = 4 \cdot (f(1) - 2 \cdot f(0))$$

We can put this in equation I, which yields the following:

$$f(2) = \frac{3}{4}\alpha_2 + 0.25 + 2 \cdot f(1) - 4 \cdot f(0)$$

By solving this for α_2, we obtain the proportion of full liars:

$$\alpha_2 = \frac{4}{3}(f(2) - 2 \cdot f(1) + 4 \cdot f(0) - 0.25)$$

B2. Additional regression analysis of the emotion shame

Furthermore, we ran an additional regression of the self-reported emotion ‘shame’, in which we included behavior and beliefs about both coin tossing tasks (see Table 2.3). Again, we found the exact same pattern, as shame was only correlated with behavior and beliefs in the non-strategic situation, but this association vanished in the competitive environment.

	Dependent Variable: *Δ shame*
Treatment (Oxytocin=1 or Placebo=0)	0.176
	(.206)
Reported outcome (competition)	-0.025
	(.256)
Believed lying rate in session j (competition)	0.605
	(.645)
Reported outcome (no competition)	0.609***
	(.223)
Believed lying rate in session j (no competition)	-1.912***
	(.630)
Constant	1.523*
	(.782)
N	120
R^2	0.0986

Table 2.3 | Full regression model on changes in the emotion ‘shame’. The self-reported level of shame is used as a dependent variable in an OLS regression. Emotion ratings after the decision in the respective coin tossing task were normalized by subtracting from reported baseline levels that were elicited immediately before the experiment. To avoid an experimenter demand effect we elicited seven other general emotions. Note that ***, **, * denote significance at the 1%, 5%, 10% level and standard errors are in parentheses.

B3. Supplemental Analysis: Accuracy of Beliefs

Furthermore, we analyzed the accuracy of beliefs, and whether subjects were good in predicting the lying rate in their respective session. We therefore calculated the absolute difference from actual lying in a session from the believed lying rate to obtain a measure for properly estimating the honesty of others. Accordingly, the absolute deviation d from the actual lying rate l in session j is denoted by $d = |l_j - b_{ij}|$, whereas b_{ij} is the guessed lying rate by subject i in session j (see Fig. 2.3).

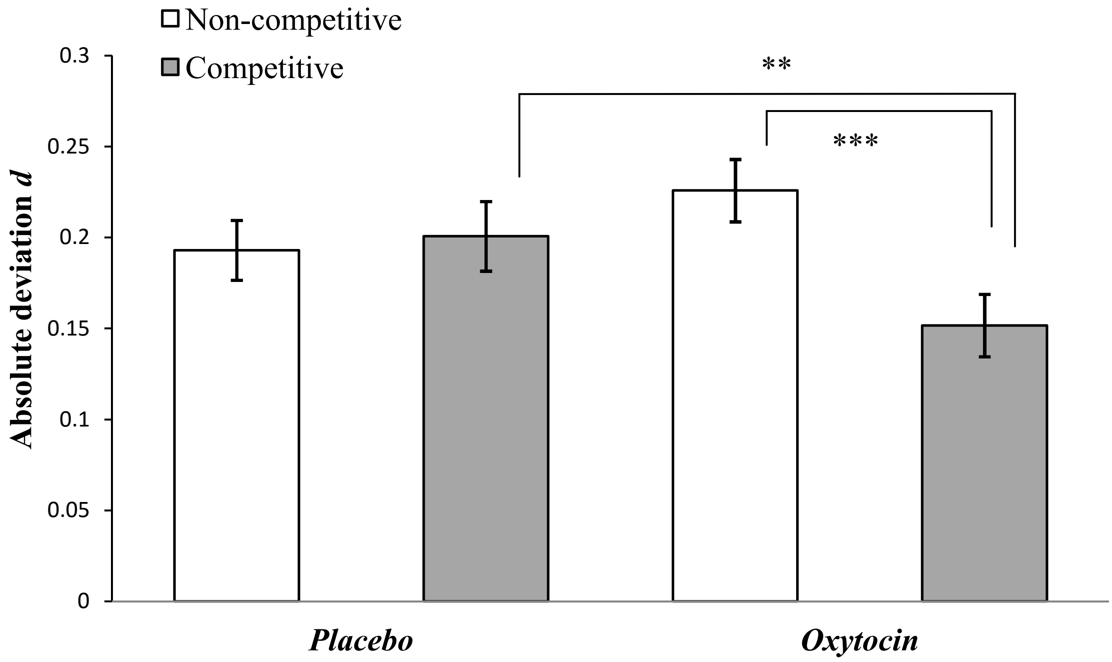


Fig. 2.3 | Absolute deviation *d* of the believed lying rate from the actual lying rate in the respective session as a function of treatment (oxytocin vs. placebo) and condition (No interaction vs. Competition).

In the competitive environment, subjects in the oxytocin group estimated the lying rate significantly more accurately (oxytocin: *mean* = 15.1% vs. placebo: *mean* = 20.0%; Mann-Whitney U-test; z = 2.219; P = 0.026). However, in the non-competitive situation when strategic considerations played no role, both treatment groups showed no statistical difference in their assessment of the lying rate in their session (oxytocin: *mean* = 22.5%, placebo: *mean* = 19.3%; Mann-Whitney U-test; z = –1.474; P = 0.140). That is, in non-strategic situations subjects under the influence of oxytocin were not better in estimating the honesty level in their session.[11]

[11] In the oxytocin group, predictions of the honesty levels were significantly better in the strategic than in the non-strategic situation (22.5% in the non-competitive vs. 15.1% in the competitive environment; Wilcoxon matched-pairs test; z = –3.114; P = 0.001).

B4. Experimental Instructions (Translated from German)

Welcome to the experiment and thank you for your participation! From now on, please do not talk to other participants and please switch off your phone!

General Procedure

The purpose of this experiment is to study decision making. You can earn money during this experiment which will be paid out to you in cash at the end of the experiment. During the experiment you and the other participants will make decisions. Your own decisions as well as those of the other participants will affect your payment according to the rules explained in the following. The duration of the experiment is 120 minutes. If you should have any questions, please raise your hand. One of the experimenters will come to you and answer your questions in private. During the experiment we do not speak of Euros, but of Guilders. Thus, your earnings during the experiment will be calculated in Guilders. At the end of the experiment your earnings in Guilders will be converted into Euros according to the following exchange rate: 3 Guilders = 1 Euro and 1 Guilder = 33 Eurocent.

For your punctual arrival you additionally receive 12 Guilders. At the end of the experiment you will be paid your total earnings from the experiment in private and in cash. While you make your decisions a clock will count down at the upper right corner of the computer screen. This countdown serves as an orientation for how much time you should need to take the decision. However, usually you can take more time to make your decision. If it is not possible to take more time than indicated, we will inform you in advance.

Anonymity

Your decisions over the course of the experiment are anonymous. You will not receive personal information about the other participants of the experiment. All data elicited will be encoded and will be treated confidentially. At the end of the experiment you will sign a confirmation which says that you have received the payment. This confirmation only serves accounting purposes for our sponsor. Our sponsor will not receive any individual data from the experiment either.

Utilities

For this experiment you are given a pen and a scratch paper. We kindly ask you to leave them at your place after the experiment.

The experiment

For each part you will receive separate instructions in written and/or verbal form. The individual parts are independent from each other, this means decisions made in one part of the experiment will not affect later parts. Your total profit will be the sum of the individual profits and the 12 Guilders you receive for punctual arrival. At the end of the experiment you will be informed about your total profit and paid individually in cash.

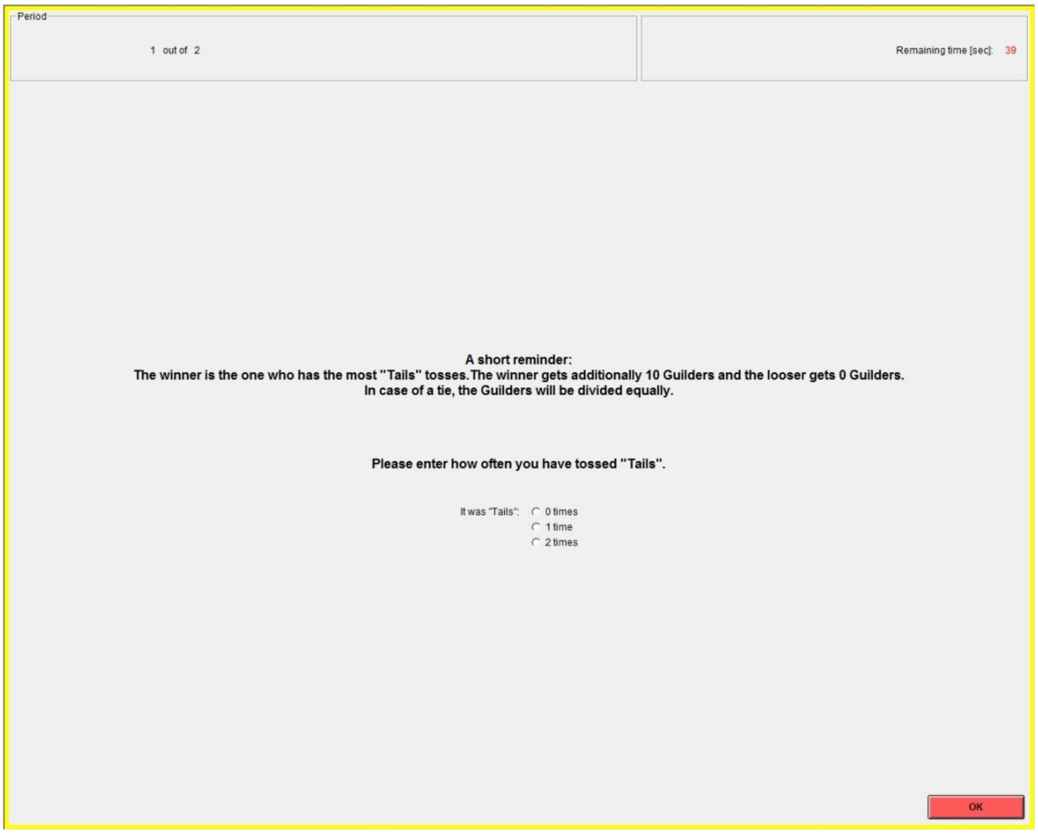
Application of nasal sprays

In the first part we ask you to apply yourself a nasal spray. Neither you nor the attending experimenters know whether you are given oxytocin or a non-effective aqueous solution (placebo). The sprays are distributed by a medical professional and the experimenter. Please refrain from application until asked to do so. All participants should apply the spray at the same time. Please do not communicate with your neighbours. If you have any private questions, please raise your hand. An experimenter will answer your questions privately.

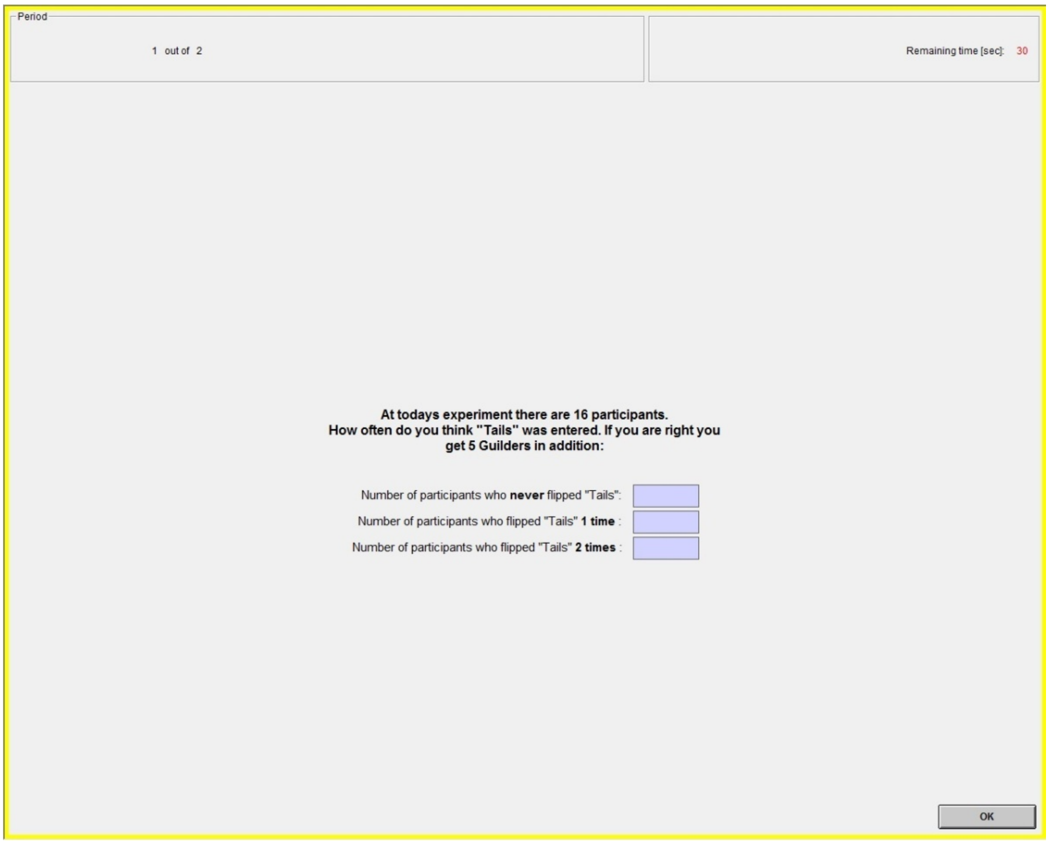
The coin tossing task

First, you get a € 1 coin from us. You will be asked to toss the coin several times and to type the outcome into the computer. Please toss the coin at your seat while sitting and make sure that neither the experimenter nor any other participant is able to see the result of your coin toss. This part consists of two sections, whereas the exact instructions of both sections will be shown on your screen.

(Screen shot of the coin tossing task)



(Screen shot of the belief elicitation task)



Chapter 3

The Adverse Effect of Oxytocin on Endogenous Group Formation*

3.1 Introduction

Decision-making in groups is a remarkable and decisive element of human societies. Humans are able to organize themselves in groups, engage in collaborative decision-making processes and are able to arrive at a binding agreement, even in the absence of an unanimous consent (Kerr & Tindale, 2004). Accordingly, numerous important decisions in the realm of economics or politics are made by small groups (Charness & Sutter, 2012; Kugler, Kausel, & Kocher, 2012). For instance, boards of directors determine the corporate strategy of a company, councils set the monetary policy of economies, and elected representatives decide on the fate of a nation. Besides the importance of participation rights and its positive effects on cooperation within societies (Frey, 1994; Sutter, Haigner, & Kocher, 2010), small groups have a striking advantage compared to individual decision-making and typically outperform individuals in intellective tasks due to a better computation of information (Laughlin & Ellis, 1986; Laughlin, Hatch, Silver, & Boh, 2006). Moreover, a growing number of empirical evidence suggests that groups exhibit a superior strategic sophistication compared to individuals (Charness & Sutter, 2012; Cooper & Kagel, 2005; Kocher & Sutter, 2005; Kugler et al., 2012). Despite the various advantages of group decision-making, the willingness to decide in groups requires a certain transfer of decision-making autonomy and a preference to deliberately expose to decisions of others. That is, a person has to trust in the ability of the group to arrive at a mutual beneficial agreement, and most importantly, to achieve a superior

---

* This chapter is based on joint work with Sandra Dehning, Andrea Jobst, Martin Kocher, Norbert Müller, Peter Zill.

outcome.[1] Thus, the ability to form (or join) a group crucially depends on the inclination to deliberately expose to decision of others.

Despite the important role of group decision-making, little or nothing is known about the neurobiological roots of the inclination to engage in collaborative decision-making. Here, we link oxytocin, an endogenous neuropeptide and hormone known for its positive effects on attachment (Donaldson & Young, 2008b), trust (Kosfeld et al., 2005) and group-cohesion (De Dreu et al., 2010; De Dreu et al., 2011), to the inclination of individuals to form groups. In particular, this work examines the effects of oxytocin on the inclination to engage in collaborative decision-making, despite the risk to bear the consequences of others' decisions. In view of the growing evidence that oxytocin plays a key role in many basic elements of social interactions that constitute a prerequisite for the maintenance of social groups like trust (Kosfeld et al., 2005), cognitive empathy (Domes et al., 2007; Shamay-Tsoory et al., 2013), positive conflict resolution (Ditzen et al., 2009) and in-group favoritism (De Dreu et al., 2011), we hypothesize that oxytocin has a positive effect on the endogenous formation of groups and the inclination to expose to decisions of others. More specifically, we hypothesize that an exogenous alteration of brain oxytocin levels has a positive effect on the willingness to transfer decision-making autonomy to a group.

3.2 Experimental Design

To test our hypothesis, we used a double-blind study design to compare group formation behavior of subjects that received a single dose of intranasal oxytocin (24 IU) with that of subjects that received a placebo, containing all ingredients except for the neuropeptide. That is, subjects were asked to solve an incentive compatible task either alone or in teams of three. Moreover, we employed a fully computerized experiment which consisted of two parts (see *Methods* for details).

About 30 minutes after participants self-administered oxytocin, we started Part One of the behavioral experiment. All participants were asked to play a cognitive and interactive task, the *p*-Beauty contest game (Nagel, 1995), for one round as individuals. That is, subjects were randomly assigned to 'units' of four and were asked to simultaneously choose a real number from the interval $I \equiv [0, 100]$. The participant whose number was closest to 2/3 times the mean of all chosen numbers in the respective unit would win a real monetary prize (12 € per Round), or an equal split in case of a tie. Accordingly, all participants had to form beliefs about their opponents' guesses, whereas the most accurate estimation of 2/3 times the mean

[1] Conversely, Knack and Keefer (1997) and Zak and Knack (2001) point out that a lack of trust in the abilities of others or the underlying decision-making process can have a detrimental effect on cooperation and collaboration in political or economic domains.

would win. Participants received feedback about their success in this part at the end of the experiment, and were informed about the subsequent second part at the end of Part One.

In Part Two, subjects were asked to choose playing the same *p*-Beauty contest game either alone or as a member of a team with three (randomly chosen) subjects. That is, three individuals and one team (consisting of three) were randomly matched to ‘units’ and played the game in the same composition for four rounds. The decision-maker (either individual or team) whose number was closest to 2/3 times the mean of all chosen numbers in the respective unit would again win a real monetary prize (12 € per Round), or an equal split in case of a tie. To keep the per capita payoff constant, each member of a winning team would receive the same amount as an individual (i.e. 12 €). Additionally, after each round all decision-makers received feedback about the winning number and their success.

After indicating their preferred role (team or individual) but before the beginning of the task in Part Two, participants were asked to answer a short questionnaire to elicit their motivation and their beliefs regarding the success of teams in this task (see appendix C1 for details).[2] Subjects who guessed the winning frequency of teams correctly would receive a bonus of 2 €. To ensure anonymity in this experiment, we employed a fully computerized setting, in which teams were able to use an interactive chat interface to discuss and agree on a particular number. They were told that in case a team was not able to agree upon a number within 4 minutes, the computer would randomly draw a number for them.[3] That is, all team members would have to agree on one particular guess, which required also the willingness to compromise, or even to subordinate to decisions of others and of course to the decision-making process itself. Accordingly, when a participant chose playing this game in a team of three, she had to be willing to expose herself to the decisions of her team members and therefore takes the risks of inferior decisions or even failed collaboration (Kocher, Strauß, & Sutter, 2006). Conversely, when a participant chose to play this game as an individual and chose to preserve her decision-making autonomy, she would sacrifice potential superior knowledge and, thus, possible superior guesses of the team (Kocher et al., 2006; Kocher & Sutter, 2005). Accordingly, from a rational cost– benefit perspective, it would only pay to opt for team decision-making if one assumes that teams would outperform individuals. Besides pure financial considerations, we were also able to examine participants’ intrinsic preferences for decision-making autonomy via a short questionnaire (see appendix C1 for details). Thus, by comparing the inclination to join a team between treatment groups, it is

[2] Since there were 12 participants in each session, and each unit consisted of 6 participants (three team members and three individuals), it was difficult to assign each participant to her preferred role. However, we were able to respect the preferences of 91.6% of all participants.

[3] This was never the case, as all subjects were able to agree on a specific number.

possible to examine oxytocin's influence on endogenous group formation and preferences for collaborative decision-making.

Moreover, the *p*-Beauty contest game constitutes an ideal tool for studying the performance of teams and individuals, as it is possible to measure the steps of reasoning that a decision maker actually applies to solve this game (Camerer et al., 2004; Coricelli & Nagel, 2009; Nagel, 1995). For instance, a player, who forms no beliefs about the guesses of other players, will pick any random number between 0 and 100 (for simplicity according to a uniform distribution), and is therefore defined as a 'step 0' thinker. Conversely, 'step 1' players will anticipate that others are 'step 0' players and will best respond by 33 (2/3 of 50).[4] Thus, a 'step 2' thinker assumes that the other players are a combination of 'step 0' and 'step 1' types and anticipates an average number between 33 and 50. According to her beliefs of the frequencies of each type, she will choose 2/3 times her estimation of the mean and so on. Accordingly, we can pursue this method with any number of iterations and can estimate how many steps a decision-maker applies to solve this task (see appendix C3 for assumptions and estimation techniques of the cognitive hierarchy model proposed by Camerer et al., 2004).[5]

3.3 Results

3.3.1 Oxytocin and decision-making preferences

According to our hypothesis participants in the oxytocin group (n = 60) should show a higher inclination to self-select into teams than those in the placebo group (n = 60). In fact, our data shows the exact opposite. Under the influence of oxytocin, participants opted considerably less often for the team. Out of 60 subjects, only 27 (45%) in the oxytocin group preferred to carry on the *p*-Beauty contest game in a team, whereas 37 of 60 subjects (62%) in the placebo group made the same decision and preferred to act in a team. The comparison of the distributions in both groups reveals a significant difference in the inclination to endogenously form teams between the oxytocin and the placebo group (see Fig. 3.1). Thus, participants given oxytocin showed a significantly decreased willingness to join a team (χ^2 = 3.348; df = 1; P = 0.067, two-sided).

[4] Note that by assumption subjects are prone to overconfidently mispredict the sophistication of others and believe that others will use a lower level of reasoning as they do (Nagel, 1995).

[5] Standard game theory suggests (for $p = ⅔$) that everyone chooses 0, the unique (Nash) equilibrium of this game Nagel (1995). According to the iterated elimination of weakly dominated strategies (IEwDS), all numbers greater than $⅔ \cdot 100$ are strictly dominated, since they would never win. However, if this is common knowledge, then everyone should eliminate all numbers greater than $⅔^2 \cdot 100$ and so on. Accordingly, all numbers greater than $p^k \cdot 100$, with $k \to \infty$, should be excluded from a player's choice set. The only remaining strategy is 0, the unique equilibrium of this game.

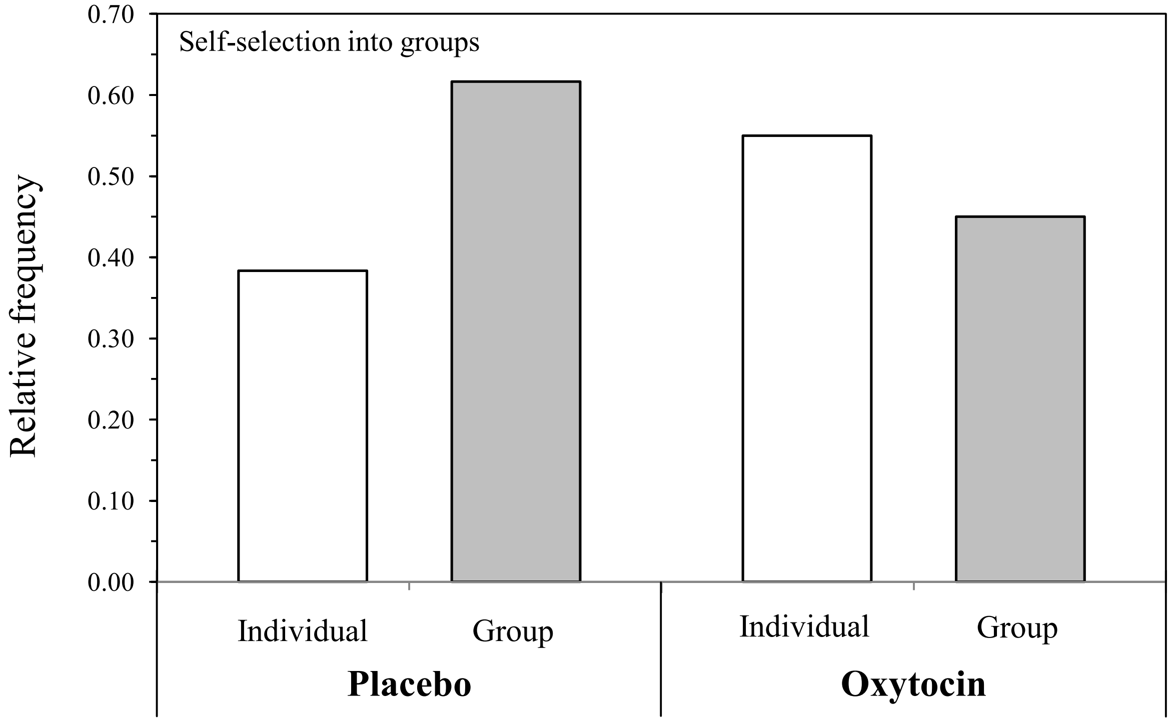


Fig. 3.1 | Self-selection into teams before the task in Part Two. Relative frequency of participants' decision to join a team (filled bars) or to continue as an individual (open bars). Subjects in the oxytocin group (N = 60) showed a significant lower inclination to endogenously form teams ($\chi^2 = 3.348$; $df = 1$; $P = 0.067$, two-sided) compared to placebo ($N = 60$).

3.3.2 Intrinsic preferences for decision-making autonomy

However, the question arises what might have caused the observed adverse effect of oxytocin on the inclination to engage in collaborative decision-making. We tested the possibility that oxytocin might have altered intrinsic preferences for decision-making autonomy, which would be independent of financial motives. That is, we examined participants' preferences for generic decision-making autonomy independent of pecuniary considerations. To measure the intrinsic preference for individual decision-making, we asked participants to rate the importance of decision-making autonomy to them on a 5-point Likert scale. According to our results the neuropeptide showed no significant influence on the preferences for autonomy (*mean* = 2.3 in placebo vs. *mean* = 2.55 in the oxytocin group; Mann-Whitney U-test; $z = -1.451$, $P = 0.1468$; two-sided test).[6]

3.3.3 Financial motives for decision-making autonomy

This leads to the possibility that the observed effect of oxytocin might be rooted in financial considerations. That is, transferring decision-making autonomy to a team would only be more

[6] Please note that preferences for autonomy were highly predictive for the decision to play the game as an individual in both treatment groups (*mean* = 2.91 for subjects preferring the role 'individual' vs. *mean* = 2.0 for subjects preferring 'team'; Mann-Whitney U-test; $z = 5.618$, $P < 0.001$; two-sided test).

beneficial if one assumes that teams would outperform individuals. In particular, subjects who were not able to comprehend the *p*-Beauty contest game would benefit the most from access to others' knowledge. Conversely, subjects with highest levels of strategic sophistication would benefit the least from team decision-making, and additionally, they would face the risk to subordinate to inferior decisions of others. To examine these motives and the effect of oxytocin on them, we analyzed subjects' performance in Part One of our experiment. In this part, subjects had the opportunity to become familiar with the task before they were asked to choose their role. That is, performance in the first part might be predictive for a subject's choice, since it is straightforward to presume that those who were not able to comprehend this task would opt for the team. Since lower guesses would indicate more steps of strategic thinking in this task and therefore a superior strategic sophistication, we linked lower guesses of a decision-maker to a better performance in this task. Hence, we regressed the preferred role (team or individual) on treatment (oxytocin or placebo) and guesses in Part One of the experiment in a Probit regression (Table 3.1). In line with the previous statistical test, the treatment had an adverse effect on the inclination to join a team. It is interesting to note that performance in the first part seemed to be not predictive for the preference for team decision-making. That is, we were not able to find evidence for the notion that subjects who did not understand the game structure would opt for the team. However, we found a remarkable interaction effect of treatment with guesses. This indicates that oxytocin had an adverse effect on subjects with low guesses (i.e. high performance), whereas this effect declined with increasing guessed numbers (i.e. low performance).

	Dependent Variable: Preferred Role (Individual = 0; Team = 1)		
Independent Variables	(1)	(2)	(3)
Treatment (Oxytocin=1 vs Placebo=0)	–.166*	–.15715*	–.1566*
	(.089)	(.090)	(.088)
Guesses in Part One		.0022	-.0014
		(.002)	(.002)
Treatment x Guesses in Part One			.0064**
			(.002)
Pseudo R^2	0.0203	0.0268	0.049

Table 3.1 | The effect of oxytocin on self-selection. Note that ***, **, * denote significance at the 1%, 5%, 10% level and standard errors are in parentheses. The preferred role, individual or team member, is the dependent variable in a Probit regression ($N = 120$ in each regression). Positive coefficients indicate positive marginal effects on probability to self-select into a team.

Fig. 3.2 shows this pattern in which the conditional marginal effect of oxytocin is depicted for three levels of guesses, i.e. mean and one standard deviation below or above average guesses. Accordingly, the influence of oxytocin is statistically insignificant on subjects with one SD above average guesses, i.e. low performers, ($z = 0.17$, $P = 0.863$, two-sided). However, the effect was greatest for subjects who guessed lower numbers, i.e. high performers ($z = -2.70$, $P = 0.007$, two-sided) and moderate for subjects with average guesses ($z = -1.72$, $P = 0.086$, two-sided). It is quite remarkable that subjects who performed above average in this task exhibited the strongest adverse effect of oxytocin on the inclination to opt for team decision-making. It seems that one important motive not to opt for a team was caused by confidence in one's own performance. As lower guesses were associated with higher levels in strategic sophistication, we concluded that the benefits for these subjects to conduct this task in a team would be the lowest. Therefore, subjects who exhibited higher levels in strategic sophistication would have the lowest gains in transferring decision-making autonomy to a team. This effect, however, was even enhanced by oxytocin, as significantly more high performers in the oxytocin group were reluctant to join a team compared to placebo.

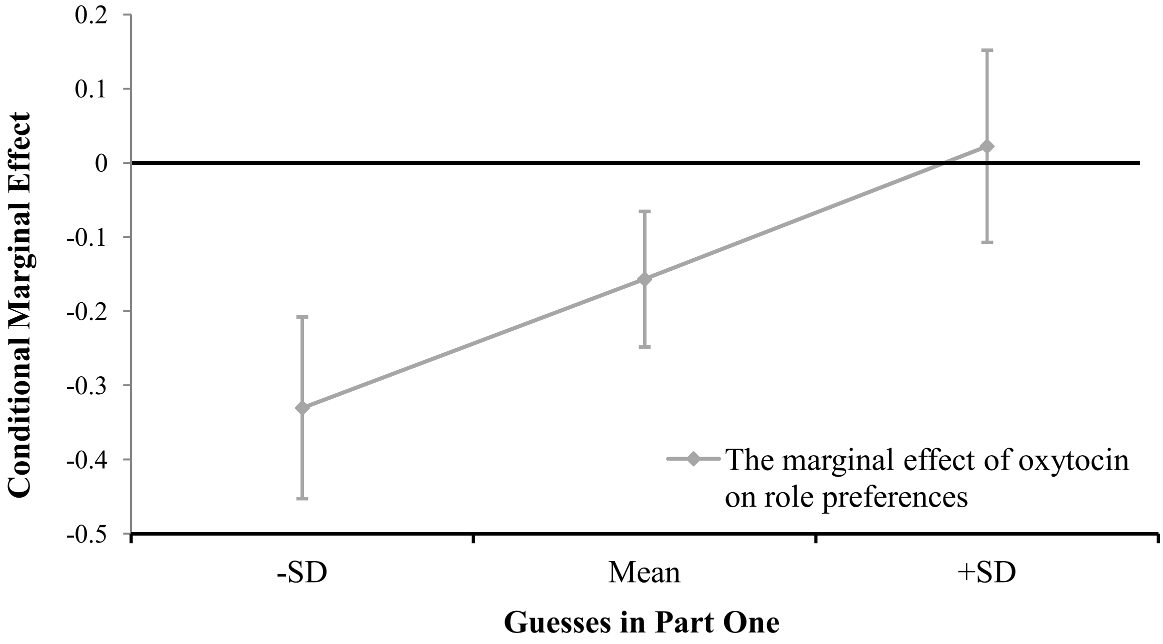


Fig. 3.2 | The conditional effect of oxytocin on the inclination to opt for team-decision-making as a function of guesses in Part One (displayed ± SE). Oxytocin exhibited a stronger adverse effect on team formation for subjects with relatively low guesses (– SD) and moderate for subjects with average guesses (Mean) in the first part of the experiment. Please note that lower guesses (– SD) indicate a higher performance in the *p*-Beauty contest game. However, the effect is insignificant for subjects with relatively high guesses (i.e. low performance).

3.3.4 Beliefs about winning chances of teams

This leads to the possibility that oxytocin might have affected the beliefs regarding the prospects of team play. Thus, we addressed this question by eliciting subjects' beliefs about the winning chances of teams. Accordingly, before the beginning of the task in Part Two, participants were asked to estimate the average winning frequency of teams in their session. If their guess was correct participants would receive an extra bonus of 2 Euros. Table 3.2 shows the results of an OLS regression with the ex ante beliefs about the winning frequencies of teams and oxytocin as dependent variables. In fact, we found a significant difference in the ex ante assessment of team performance. According to our results, subjects in the oxytocin group were significantly more pessimistic about the winning chances of a team, since the average guessed winning frequency was 42.2% in the placebo group and 35.6% in the oxytocin group ($P < 0.05$ in model 1).[7] Thus, it seems that oxytocin negatively influenced subject's beliefs regarding the success of teams. Remarkably, this adverse effect of the neuropeptide was still (weakly) significant ($P < 0.1$), if controlled for the actually chosen role of a subject (team vs. individual). This indicates that subjects given oxytocin were more pessimistic about the

[7] Remarkably, in both treatment groups subjects anticipated that teams would significantly outperform compared to their statistical chance of winning of 25% (Wilcoxon signed-rank test; $z = 5.420$, $P < 0.001$ in the placebo group; $z = 4.288$, $P < 0.001$ in the oxytocin group, two-sided).

success of teams in this task, irrespective of their preferred role. Furthermore, we found that subjects, who preferred to conduct this task in a team, would also be more optimistic about the winning chances of their team ($P < 0.05$). This is not surprising, as subjects were incentivized to win this task and therefore would choose the role they assign the highest chances to win this task. We conclude that subjects given oxytocin were more pessimistic about the success of teams, and therefore showed a significantly lower inclination to join a team.

	Dependent Variable: *Believed winning frequency of teams*	
	(1)	(2)
Treatment (Oxytocin=1 or Placebo=0)	–0.066**	–0.055*
	(0.033)	(0.033)
Preferred Role (Team=1 or Individual=0)		0.068**
		(0.033)
Constant	0.422***	0.380***
	(0.023)	(0.031)
N	120	120
R^2	0.032	0.066

Table 3.2 | The effect of oxytocin on the believed performance of teams. The believed winning frequency of teams is the dependent variable in an OLS regression. Negative coefficients indicate a more pessimistic ex ante assessment of team performance in the *p*-Beauty contest game. Note that ***, **, * denote significance at the 1%, 5%, 10% level.

To examine the causes of the more pessimistic view in the oxytocin group, it is necessary to analyze the ex ante beliefs of team performance in more detail. In particular, we addressed the question whether the ex ante beliefs regarding the performance of teams were justified, i.e. more realistic, in the oxytocin group. Therefore, we computed the absolute difference from the actual winning frequency from the believed winning frequency to obtain a measure for properly estimating the success of teams. Accordingly, the absolute deviation d from the actual success rate s is denoted by $d = |s - b_i|$, whereas b_i is the guessed (or believed) success rate by subject i. Remarkably, we found that subjects in the placebo group were not good in estimating the success rate, as they significantly overestimated the winning frequency of teams (Wilcoxon matched-pairs test; $z = 2.576$; $P = 0.01$). Conversely, subjects in the oxytocin group were surprisingly good in assessing the success of teams as their estimation of the success rate was statistically not different from the actual success rate of 32.5% (Wilcoxon matched-pairs test; $z = 0.646$; $P = 0.518$). According to our results, the assessment of teams was biased in the placebo group, but not in the oxytocin group, as their estimates were statistically not distinguishable from the actual success rate of teams. We conclude that subjects in the oxytocin group were not more pessimistic, but rather more realistic about the winning chances of teams compared to placebo.

Our findings indicate that subjects given oxytocin were better in assessing the risks and benefits of team decision-making. This would also explain why oxytocin had a stronger effect on subjects with higher strategic sophistication. As the benefits of team decision-making would be the lowest for subjects with high strategic sophistication, since their gains from the access to knowledge of others would be little. As oxytocin induced a more pessimistic (but realistic) assessment of the benefits of collaborative decision-making, subjects with higher sophistication might have perceived individual decision-making as more beneficial. Thus, beliefs about the performance of teams could have constituted a mediator between oxytocin and preferences for team decision-making (Baron & Kenny, 1986), since our previous findings indicate significant correlations of all three variables with each other (i.e. oxytocin, beliefs and preferred role).[8] To examine the mediating effect of participants' beliefs, we ran a further Probit regression with the chosen role as a dependent variable and oxytocin as well as beliefs as independent variables (see appendix C2 for details). The regression reveals that beliefs were still a good predictor for the chosen role ($\beta = 0.524$, SE = 0.257, $P = 0.042$), however, the effect of oxytocin declined and failed significance ($\beta = -0.137$, SE = 0.092, $P = 0.135$). Thus, we conclude that the effect of oxytocin on the preferences to join a team was mediated by beliefs, which made up about 19.4% of the total effect (see appendix C2 for details).

3.3.5 Performance of teams in the *p*-Beauty contest game

Furthermore, we examined the actual performance of subjects in their finally assigned roles. This would give us an indication whether subjects' ex ante preferred roles were optimal from an ex post perspective. Therefore, we analyzed the guessed numbers of all four rounds in Part Two of the experiment. Table 3.3 illustrates the results of a Random effects regression with guessed numbers g_{it} of decision-maker i in round t as a dependent variable. We found no significant effect of the chosen role on guesses. However, subjects in the oxytocin group chose significantly lower numbers compared to placebo. Since lower guesses would indicate a better comprehension of this strategic task and therefore a superior strategic sophistication,

[8] Please note that the variable oxytocin was completely exogenous, since we randomly assigned subjects to both conditions. Thus, a correlation with oxytocin also indicates causation.

we were able to conclude that oxytocin had a significant positive impact on the strategic sophistication of participants. Remarkably, this effect is independent of the type of the decision-maker (team vs. individual). Our results suggest that teams were not better than individuals in this task and therefore were not able to provide access to superior knowledge. Accordingly, from an ex post view subjects were not able to improve their success in this task by opting for team decision-making.

	Dependent Variable: *Guessed number g_{it} of decision-maker i in round t*	
	(1)	(2)
Treatment (Oxytocin=1 or Placebo=0)	–4.197*	–4.197*
	(2.340)	(2.348)
Actual Role (Team=1 or Individual=0)		–0.085
		(1.629)
Round t		–7.558***
		(0.538)
Constant	16.35***	35.27***
	(1.53)	(2.255)
N	320	320
R^2	0.056	0.056

Table 3.3 | The effect of oxytocin on performance in Part Two. Guessed number g_{it} of decision-maker i in round t is the dependent variable in a Random effects estimation. Please note that in our experiment three individuals and one team are randomly matched to 'units' and play the game in the same composition for four rounds. Thus, standard errors in parentheses are clustered on unit level. Negative coefficients indicate lower guesses and therefore a higher strategic sophistication. Note that ***, **, * denote significance at the 1%, 5%, 10% level.

3.3.6 Oxytocin and performance in the *p*-Beauty contest game

Remarkably, oxytocin had a positive effect on performance in the *p*-Beauty contest game. However, to compute an exact measure of subjects' performance, we estimated the applied steps of strategic thinking for each treatment group. The computation of the exact steps was obtained by analyzing the first part of the experiment, in which each guessed number per subject was an independent observation. Accordingly, we employed the cognitive hierarchy model proposed by Camerer et al. (2004) for the estimation, which constitutes a useful tool to describe behavior in this game (see appendix C3 for more details). According to the results (see Table 3.4), participants under the influence of oxytocin applied 1.53 steps, whereas in the placebo treatment subjects reached a level of 0.86 (post-estimation Wald test; z = 1.374; P = 0.0844, one-sided, N = 60 in the oxytocin and N = 60 in the placebo group). That is, the exogenous alteration of brain oxytocin levels resulted in a substantial enhancement of strategic reasoning, as the applied steps almost doubled. That is, under the influence of oxytocin subjects showed an increased ability to rationalize possible strategies of their co-players.

	Estimate of iterations τ	
	Placebo group	*Oxytocin group*
Steps of thinking in Part One	***.861***	***1.528***
	(.333)	*(.388)*

Table 3.4 | Steps of thinking in Part One. Parameter τ was estimated by GMM estimation techniques with N = 60 for each treatment group. Robust standard errors are in parentheses.

3.4 Discussion

The ability to engage in group decision-making is one of the most decisive elements of human societies. Group decision-making plays a key role in a variety of social, political or economic domains and constitutes therefore a key element for social interactions. Here we examined the effects of oxytocin on the inclination to endogenously form groups. We provide evidence that an exogenous alteration of brain oxytocin levels causes a significant adverse effect on group formation, which contradicts standard hypotheses regarding the role of oxytocin in human behavior (Bethlehem et al., 2014; MacDonald & MacDonald, 2010). That is, subject given oxytocin exhibited a significantly lower propensity to opt for solving a strategic task, the *p*-Beauty contest game, in a team compared to placebo. Replicating past results, about 62% opted for the team decision-making in the placebo group (Kocher et al., 2006), whereas only 45% were willing to engage in collaborative decision-making in the oxytocin group. Our results indicate that this effect was most pronounced for subjects with relatively high strategic sophistication, as they showed a significantly lower propensity to opt for team decision-making due to our treatment. Furthermore, we show that the observed treatment effect of oxytocin was mainly driven by more realistic beliefs regarding the success of teams in this task. In particular, our results indicate that subjects given oxytocin were better in assessing the winning frequency of teams, as their estimates were not statistically different from their actual success rate. Conversely, subjects in the placebo group significantly over-estimated the chances of teams in this task, which explains the significant difference in the inclination to engage in collaborative decision-making between both treatment groups.

Moreover, our results suggest that an intranasal application of oxytocin also had a positive effect on strategic reasoning. That is, subjects given oxytocin chose significantly lower numbers in the *p*-Beauty contest game, irrespective of the actual role (i.e. team or individual). These findings contribute to previous studies that demonstrated the important role of oxytocinergic systems in cognitive empathy, the ability to put oneself emotionally in the shoes of another person (Domes et al., 2007; Guastella et al., 2010; Rodrigues, Saslow, Garcia, John, & Keltner, 2009). Here, we show that oxytocin has a similar positive effect in strategic situations, as subjects given oxytocin showed an enhanced ability to take the strategic perspective of others and to strategically mentalize. In view of that, it is not surprising that a dysfunction of the oxytocinergic system is related to disorders, which are characterized by impairments in social cogniton and mentalizing abilities, like autism spectrum disorder (Jacob et al., 2007; Modahl et al., 1998) and schizophrenia (Pedersen et al., 2011). Consequently, our results give rise to the notion that the cognitive processes related to strategic reasoning (Bhatt & Camerer, 2005; Coricelli & Nagel, 2009; Hampton, Bossaerts, & O'Doherty, 2008) might also be related to the oxytocinergic system.

However, our results do not support the notion that teams in general perform better than individuals in intellective tasks (Laughlin & Ellis, 1986; Laughlin et al., 2006) or exhibit a superior strategic sophistication compared to individuals (Charness & Sutter, 2012; Cooper & Kagel, 2005; Kocher & Sutter, 2005; Kugler et al., 2012). In fact, teams exhibited the same performance as individuals, even if controlled for our treatment. However, we argue that our design might have biased the performance of teams, as subjects were free to choose their preferred role. In line with this notion, we found an adverse self-selection process due to our treatment, since the observed negative effect of oxytocin on group formation was even more pronounced for subjects with relatively high strategic sophistication. This finding indicates that subjects in the oxytocin group with high levels of strategic sophistication anticipated a risk to subordinate to inferior decisions of others and might have viewed team decision-making as an inferior option.[9] Therefore, fewer participants with high strategic sophistication joined teams, which ultimately led to a decrease in their performance. Combined with the finding that oxytocin enhanced strategic sophistication, we argue that subjects in the oxytocin group might have anticipated that less participants with high strategic sophistication would join teams. This would also explain the more pessimistic assessment of team’s winning

[9] In particular, team decision-making could bear the risk of disagreements or bad compromises, since each team member had to agree on a specific number.

chances compared to placebo. That is, subjects might have anticipated this adverse selection process and therefore might have assigned a lower winning chance to teams. From a rational cost– benefit perspective, it would not have paid for engaging in teams if one assumed that only low performers would opt for team decision-making. As a result, high performers would anticipate this adverse self-selection process and would prefer individual decision-making in the first place. [10] Accordingly, our results indicate the importance of the causes and consequences of self-selection processes on team performance, as teams may attract more high performers (Hamilton, Nickerson, & Owan, 2003) or low performers (Kocher et al., 2006) depending on the underlying incentive structures. Here, we show that a chemically induced change in beliefs about the success of teams could have an adverse effect on their composition, and ultimately on their performance in a competitive task.

Previous empirical studies were able to show that humans predominantly prefer to engage in social interactions and organize in teams or groups (van Beest & Williams, 2006), since small groups provide benefits for the evolutionary fitness and survival of an individual (Gruter & Masters, 1986). Accordingly, team decision-making can contribute to mutual financial success (Feri, Irlenbusch, & Sutter, 2010; Hamilton et al., 2003) and provide access to superior knowledge or computational power (Laughlin & Ellis, 1986; Laughlin et al., 2006). However, this work demonstrates that collaboration within teams has to pay for the individual, especially in the case when individuals exhibit heterogeneous aptitudes. Otherwise it might be hard to endogenously form groups if the underlying incentive structure would attract less high performers, which might have a detrimental effect on the performance of teams. This is certainly of importance for policy makers, companies and organizations that have to consider the consequences of self-selection into team decision-making.

3.5 Conclusion

In view of previous results regarding the behavioral effects of oxytocin (Bartz et al., 2011; Bethlehem et al., 2014), our results seem to be questioning current hypotheses regarding the underlying mechanisms of oxytocin's efficacy. In particular, we argue that our results seem to be hardly compatible with the notion that oxytocin enhances social cohesion by promoting pro-sociality (MacDonald & MacDonald, 2010), in-group favoritism (De Dreu et al., 2011) and in-group cooperation (De Dreu et al., 2010). This work gives rather rise to the notion of a more general underlying effect of oxytocin, which might be rooted in an enhanced ability of strategic perspective taking and mentalizing, as subjects given oxytocin showed a significantly enhanced performance in the strategic task. Accordingly, additional research on

[10] This argument corresponds to the notion of adverse selection in markets for 'lemons' proposed by Akerlof (1970).

this issue is indispensable to infer general conclusions on the effects of oxytocin on human behavior, which might ultimately lead to a theoretical approach or hypothesis that is able to explain likewise all well-documented effects of oxytocin.

3.6 Methods

3.6.1 Participants

A total of 120 male subjects participated in the experiments (all students from different universities in Munich). Participants were between 18 and 35 years old (mean age ± s.d., 23.4 ± 3.2 years). Exclusion criteria for participation were any kind of endocrine or neurological disorder, any kind of current or life time diagnosis of psychiatric disorder, life time participation in psychotherapy, any kind of medication, and drug or alcohol abuse. Subjects were instructed to abstain from alcohol, smoking and caffeine for 24 h, and from food for 2 h before the experiment. At the time of recruitment participants were informed that the experiment is about the influence of oxytocin on economic decision making. The study was approved by the ethics committee of the medical department at the University of Munich. All participants provided written informed consent before participation.

3.6.2 Substance administration

We used a double-blind, placebo-controlled study design to compare behavior in a group of subjects that received a single intranasal dose of 24 IU oxytocin (Syntocinon-Spray, Sigma-Tau, 3 puffs per nostril, each with 4 IU of oxytocin) with that of subjects in a control group that received a placebo. The placebo spray contained all same inactive ingredients except for the oxytocin. The economic part of the experiment started 30 minutes after the administration of the substances, as central nervous oxytocin levels reach their plateau around 30 minutes after substance administration (Born et al., 2002; Gossen et al., 2012).

3.6.3 Behavioral experiment and questionnaires

We employed a fully computerized and anonymous experimental design to ensure that participants could not sense or use any social cues of their opponents other than provided by us, since there is evidence that subjects given oxytocin react stronger to social information (Bartz et al., 2011; Declerck et al., 2010; Declerck et al., 2013). Subjects were separately seated in cubicles, whereas all decisions were made privately on computer screens. First, subjects received neutrally written instructions, which we read aloud (see appendix C4 for details). The experiment consisted of several parts which were independent of each other. To proceed to the next part of the experiment, the previous part had to be accomplished by all participants. 30 minutes after substance administration we started with the behavioral experiment, which consisted of two parts.

In Part One, all subjects played exactly one round of the *p*-Beauty contest game in units of four with each other. They were asked to simultaneously choose a real number from the interval $I \equiv [0, 100]$. The winner is the decision-maker whose number is closest to 2/3 of the mean of all chosen numbers and receives a real monetary prize (12 € per Round), or an equal split in case of a tie. Subjects received no feedback about the outcome of the first part until the end of the experiment and were informed about the subsequent second part at the end of Part One.

In Part Two, subjects were asked to choose playing the same *p*-Beauty contest game either alone or as a member of a team with three (randomly chosen) subjects. After indicating their preferred role (team or individual), participants were asked to answer a short questionnaire to elicit their motivation and their beliefs regarding the success of teams in this task (see appendix C1 for more details). Subjects who guessed the winning frequency of teams correctly would receive an extra bonus of 2 €. Then, according to the indicated preferred roles, we randomly assigned three individuals and one team (consisting of three) to 'units' and asked them to play the game in the same composition for four rounds. The decision-maker (either individual or team) whose number was closest to 2/3 times the mean of all chosen numbers in the respective unit would again win a real monetary prize (12 € per Round), or an equal split in case of a tie. To keep the per capita payoff constant, each member of a winning team would receive the same amount as an individual (i.e. 12 €). To ensure anonymity in this experiment, we employed a fully computerized setting, in which teams were able to use an interactive chat interface to discuss and agree on a particular number. Teams had 4 minutes to agree upon a number, otherwise the computer would randomly draw a number for them.[11] Furthermore, all decision-makers were only informed about the target number in their unit (i.e. 2/3 times the mean) and about a decision-maker's own success in the respective round. No information was provided about their co-players chosen numbers or their success. In the end of the experiment all participants received a show up fee of 12 € and their payoff corresponding to their success in the experiment.

[11] This never happened, as all groups were able to agree on a number within the time frame.

Appendix C

C1. Short Questionnaire and Beliefs Elicitation

Immediately after subjects indicated their preferred role, they were asked to rate the following eight statements on a five point Likert scale between 1 (not at all) and 5 (totally agree).

Reasons for your decision:

1. This way I will earn more money.
2. I want to bear the responsibility for my own decisions.
3. I don´t want to discuss.
4. It is hard for me to trust others.
5. Decision-autonomy is important for me.
6. Small teams are better decision-makers than individuals.
7. It is boring to decide as an individual.
8. It is hard for me to explain my thoughts to others.

None of the rated items showed significantly different ratings between treatment groups (Mann-Whitney U-test; P > 0.1; two-sided test).

Beliefs Elicitation

Immediately after the questionnaire, subjects were asked to estimate the winning frequency of both teams in their respective session. For a correct prediction, they would receive 5 MU, which is equivalent to 2 €. The question on the computer screen was as follows:

"If both, individuals and teams, are equally good decision-makers, then their winning chance should be 25%. In your opinion, what will be the winning frequency (in percent) of teams in this session? If you are right you will receive an extra bonus of 5 MU."

C2. Mediation analysis

	Dependent Variable: Preferred Role (Individual = 0; Team = 1)		
Independent Variables	(1)	(2)	(3)
Treatment (Oxytocin=1 vs Placebo=0)	–.166*		–.137
	(.089)		(.092)
Believed winning frequency of teams		.580**	.524**
		(.252)	(.257)
Pseudo R^2	0.0203	0.0323	0.0456

Table 3.5 | The effect of oxytocin and beliefs on self-selection into teams. Note that ***, **, * denote significance at the 1%, 5%, 10% level and standard errors are in parentheses. The preferred role, individual or team member, is the dependent variable in a Probit regression. Positive coefficients indicate a positive marginal effect on the probability to self-select into a team.

We computed the causal mediation effect with the method proposed by Hicks and Tingley (2011), which is able to compute casual mediation effects for binary and continuous dependent variables (i.e. suitable for OLS, Probit and Logit regressions). For a continuous mediator variable and a continuous outcome variable, the results are identical to the method proposed by Baron and Kenny (1986).

C3. The Assumptions of the Cognitive Hierarchy Model

The cognitive hierarchy model (Camerer et al., 2004) assumes that each player is a step-k player, and has accurate beliefs about the proportions of players doing h steps of thinking. This proportion of players doing h steps is given by:

$$g_k(h) = f(h) / \sum_{l=0}^{k-1} f(l), \forall h < k$$

Note that everyone assumes that there are no players applying higher steps of reasoning as they do, such that $g_k(h) = 0$, $\forall\ h \geq k$. Furthermore, step-0 players are not thinking strategically and, for simplicity, equally randomize over all strategies.

Then, a step-k player best responds to all other players according to his beliefs about the proportions of all lower types, given by $f(0), f(1), f(2), \dots f(k-1)$.

To obtain plausible properties of the distribution $f(k)$, it is necessary to define further requirements of it. For instance, the distribution of $f(k)$ should be discrete, since the thinking steps k are integers. And of course, the relative proportion $f(k) / f(k-1)$ should decline with k, as more steps of thinking need more cognitive capacity, which is by assumption a limited resource. If one assumes that the decline of $f(k) / f(k-1)$ is proportional to $1/k$, then $f(k)$ is the Poisson distribution, and is given by:

$$f(k) = \frac{e^{-\tau}\tau^k}{k!}$$

Note that τ represents the mean of employed thinking steps and its variance. This gives the possibility to estimate the average steps of thinking τ by using GMM estimation techniques.

GMM Estimation of the reasoning parameter τ

As step-0 players randomize over the whole interval of [0,100], step-1 players will best-respond to them by choosing 33 (for $p = ⅔$). However, step-2 players will best-respond to a linear combination of step-1 and step-0 players. That is, step-2 players will weight 33 and 50 by the frequencies of step-1 and step-0 players in the population. Thus, step-2 players will best respond to $(50 \cdot f(0) + 33 \cdot f(1)) / (f(0) + f(1))$. As $f(k)$ is Poisson distributed, a step-2 player will best respond to $(50 + 33 \cdot \tau) / (1 + \tau)$, and so on.

Consequently, we were able to estimate τ separately for each treatment condition (oxytocin and placebo) by using the generalized method of moments. That is, for each treatment group we minimized the difference between the predicted mean and the actual mean of guesses.

C4. Experimental Instructions *(Translated from German)*

Welcome to the experiment and thank you for your participation! From now on, please do not talk to other participants and please switch off your phone!

General Procedure

The purpose of this experiment is to study decision making. You can earn money during this experiment which will be paid out to you in cash at the end of the experiment. During the experiment you and the other participants will make decisions. Your own decisions as well as those of the other participants will affect your payment according to the rules explained in the following. The duration of the experiment is 150 minutes. If you should have any questions, please raise your hand. One of the experimenters will come to you and answer your questions in private. During the experiment we do not speak of Euros, but of Points. Thus, your earnings during the experiment will be calculated in Points. At the end of the experiment your earnings in Points will be converted into Euros according to the following exchange rate:

2.5 Points = 1 Euro

1 Point = 40 Eurocent.

For your punctual arrival you additionally receive 30 Points. At the end of the experiment you will be paid your total earnings from the experiment in private and in cash. While you make your decisions a clock will count down at the upper right corner of the computer screen. This countdown serves as an orientation for how much time you should need to take the decision. However, usually you can take more time to make your decision. If it is not possible to take more time than indicated, we will inform you in advance.

Anonymity

Your decisions over the course of the experiment are anonymous. You will not receive personal information about the other participants of the experiment. All data elicited will be encoded and will be treated confidentially. At the end of the experiment you will sign a confirmation which says that you have received the payment. This confirmation only serves accounting purposes for our sponsor. Our sponsor will not receive any individual data from the experiment either.

Utilities

For this experiment you are given a pen and a scratch paper. We kindly ask you to leave them at your place after the experiment.

The experiment

For each part you will receive separate instructions in written and/or verbal form. The individual parts are independent from each other; this means decisions made in one part of the experiment will not affect later parts. Your total profit will be the sum of the individual profits and the 30 Points you receive for punctual arrival. At the end of the experiment you will be informed about your total profit and paid individually in cash.

Application of nasal sprays

In the first part we ask you to apply yourself a nasal spray. Neither you nor the attending experimenters know whether you are given oxytocin or a non-effective aqueous solution (placebo). The sprays are distributed by a medical professional and the experimenter. Please refrain from application until asked to do so. All participants should apply the spray at the same time. Please do not communicate with your neighbors. If you have any private questions, please raise your hand. An experimenter will answer your questions privately.

The Task

The experiment consists of two distinct parts. At first, you will get some instructions for Part One. Your decisions in Part One have no consequences on Part Two. You will receive the instructions for Part Two after Part One.

Part One:

First, three units each consisting of four people ("decision-makers") are formed. The allocation will be done randomly. Now, each member of the unit has to choose a number out of the interval [0, 100] by entering the number on the respective screen on the computer. Choosing 0 or 100 is also possible.

The decision-maker whose number is closest to 2/3 times the mean of all chosen numbers in the group wins. (i.e. target number= 2/3* the average of all chosen numbers in your group)

In mathematical terms, the target number is calculated as follows:

$$\bar{x} = \frac{2}{3} \frac{\sum_{i=1}^{n} x_i}{n}$$

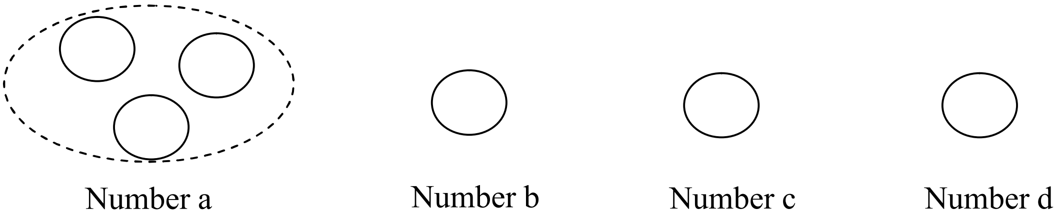
You win if your chosen number is closest to the target number. The winner receives 30 points. In case of a tie (that is several people whose chosen numbers are same and have won), the profit will be evenly distributed, i.e. divided by the number of winners. At the end of the experiment, you will get feedback on your success in this task.

Using the calculator:

In order to make calculations, you can click on the symbol of the calculator at the bottom of the screen which opens the Windows-Calculator. (Please keep the 'order-of-operation-rule' in mind!)

Part Two: (*will be distributed* _after_ *completion of Part One*)

Part Two consists of four periods with the same task as in Part One. Again, units out of four decision-makers will be formed. The composition of the unit remains constant over the four periods. However, this time one unit consists of **three individuals** and a **three-member group**. In each period, the group of three has to choose a common number; and also each individual has to choose a number in each period (see chart).



Again, the decision-maker (group or individual) whose number is closest to 2/3 times the mean of all chosen numbers wins. The target number is calculated as follows:

$$\bar{x} = \frac{2}{3}\frac{\sum_{i=1}^{n} x_i}{n}$$

After each period, you will get feedback about the target number in this period and whether you have won or not. In each period of the game, you can win 30 points. If a group wins, every single member of the group receives 30 points in this period. If an individual wins, she also receives 30 points. If more than one decision-maker wins, the earnings will be split equally among all winning decision-makers. (Example: If one individual and one group win in a period, the individual would receive 15 points and every single group member would receive 15 points.)

Before we start Part two, you will have to decide whether you want to play this game in all four periods **as an individual** or in **a group** consisting of three persons.

If you decide to conduct this task as an individual, you will have to pick a number in each period as an individual.

If you decide to conduct this task as a group, you will have to jointly agree on one particular number in each period. Please note that each group member's identity remains anonymous, i.e. none of the group members know with whom they are matched. Furthermore, the group composition remains constant over all four periods. That means, if you decided to play this task as a group member or individual, you will stay in this role over all four periods. For instance, it is not possible to opt for the group in the first two periods and then exit the group and go on as an individual.

You can communicate with the other members of your group via messages in a real-time chat. The communication rules within the chat are explained below. Only the group members can communicate within their respective group; communication between individuals or between individuals and groups is not possible.

Since there are 12 persons in this session, we will form two units that consist of one group and three individual decision-makers. We will allocate you randomly to the units according to your indicated preference (group or individual). We will try to respect the preference of as many participants as possible. However, whose preference may not be fulfilled will be randomly determined.

Communication within the group in each period

The content of your communication is up to you, but it is not allowed to reveal any personal information of you or others. These are e.g. name, age, address, your subject of study or anything else that could reveal your identity (like your seat number or place). Furthermore, it is not allowed to arrange any rewards or punishments within your group. In addition, any offensive or aggressive messages are prohibited. Anyone who violates the communication rules will be excluded from the experiment and will not receive any payment in the experiment. Each group member can send as many messages as she wants to both other group members. Messages in the chat are anonymous, i.e. in each period each subjects gets a random Chat-ID. Your Chat-ID will automatically appear on the screen. Furthermore, every message from you will appear on the screens of the other two members of your group. Selective messages to a single group member are not possible.

The group discussion automatically ends after **a maximum of 4 minutes** in each period. Please mind the clock in the upper right of the screen. After the expiration of the time limit you can enter a number. When you have completed your entry, click OK. **All three inputs of all three group members have to be identical.** Otherwise, the decision is invalid, and you cannot win in that period. In this case, the computer randomly selects a number, which is relevant for the other three individuals in your unit.

Using the calculator:

In order to make calculations, you can click on the symbol of the calculator at the bottom of the screen which opens the Windows-Calculator. (Please keep the 'order-of-operation-rule' in mind!)

If you have any questions or something remained unclear, please raise your hand. One of the experimenter will come to you and answer your questions in private.

Chapter 4

Oxytocin promotes strategic reasoning

4.1 Introduction

To understand emotions, intentions and desires of others is an inherent ability of humans, and indispensable to sustain social, political or economic interactions (Singer & Fehr, 2005). Cooperation (McCabe, Houser, Ryan, Smith, & Trouard, 2001), coordination and even competition (Hampton et al., 2008) between individuals require some degree of perspective-taking, on an emotional or cognitive level. Choosing your actions without considering others' behavior is prone to mis-coordination or exploitation by others. A famous example of this was given by the influential economist John Maynard Keynes. He pointed out that stock market speculation is very similar to beauty contests in newspapers "in which the competitors have to pick out the 6 prettiest faces from a hundred photographs, [and] the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole (Keynes, 2006)." According to his gedankenexperiment, speculative trading refers to strategic perspective-taking in a competitive environment, in which speculators try to pick assets they think that others would buy (first order belief), or buy what others think that they would buy (second order belief) and so on. Obviously, low-level reasoning (e.g. self-referential thinking) is prone to exploitation by speculators who apply higher levels of reasoning, since they are able to take low-level thinking into account and accordingly are able to sell or buy before too many others do (Coricelli & Nagel, 2009). Psychologists define these cognitive processes as the theory of mind (ToM) or mentalizing, as a person has to infer the thoughts, intentions, emotions and beliefs of others in order to correctly predict their behavior (Amodio & Frith, 2006).

Although strategic reasoning plays a crucial role in various different domains of human social interactions, there is still scarce evidence regarding the modulating role of hormones on strategic perspective-taking, i.e. the ability to understand the strategies and payoffs of others and to correctly predict their future actions in a purely strategic setting. In this study, I hypothesize that oxytocin, a neuropeptide produced in the hypothalamus and released both into brain and bloodstream (Donaldson & Young, 2008a), has a positive effect on strategic thinking. Several studies provide evidence that oxytocin improves the ability to recognize emotions in facial expressions (Domes, Heinrichs, Michel, Berger, & Herpertz, 2007; Guastella et al., 2010; Rodrigues, Saslow, Garcia, John, & Keltner, 2009). Moreover, a dysfunction of the oxytocinergic system is related to autism (Andari et al., 2010; Jacob et al., 2007; Modahl et al., 1998), a disorder characterized by impairments in social interaction, social cognition and perspective-taking. Besides its effects on cognitive empathy and pro-sociality (Kosfeld, Heinrichs, Zak, Fischbacher, & Fehr, 2005; Zak, Stanton, Ahmadi, & Brosnan, 2007), nothing or little is known about the role of oxytocin in strategic perspective-taking, i.e. the ability to understand the strategies and payoffs of others and to correctly predict their future actions in a purely strategic setting. However, due to its positive effects on the emotional part of mentalizing, it is straightforward to presume that an exogenous variation of brain oxytocin levels has a positive effect on strategic reasoning.

4.2 Experimental design

Oxytocin's role in strategic thinking was examined by employing a purely cognitive and interactive task that is commonly used to measure strategic sophistication – the p-Beauty contest game. In this task, all participants have to pick a real number from the interval $I \equiv [0, 100]$. The winner, whose number is closest to 2/3 times the mean of all chosen numbers, wins (see *Methods*). The p-Beauty contest game is an ideal tool for studying the cognitive processes of iterated steps of reasoning (Nagel, 1995) and provides a measure of the actually applied steps to solve this game (Camerer et al., 2004) (see Fig. 4.1). For instance, a player, who forms no beliefs about the guesses of other players, will pick any random number between 0 and 100 (for simplicity according to a uniform distribution). Since this player has no beliefs about the opponent's guesses, such behavior is commonly defined as level 0 thinking. However, a level 1 player will anticipate that other players will pick a random number and expect therefore an average of 50. Thus, the best response of a level 1 thinker $L(1)$ is 33 ($= 2/3 \cdot 50$). However, a level 2 thinker $L(2)$ assumes that the other players are level 1 types and chooses 22 ($= 2/3^2 \cdot 50$) and so on. Accordingly, one can pursue this method with any number of iterations and can estimate how many steps a decision-maker applies to solve this task. It is important to note that by assumption subjects are prone to overconfidently mispredict the sophistication of others and believe that others will use a lower level of reasoning as they do (Camerer et al., 2004; Nagel, 1995).

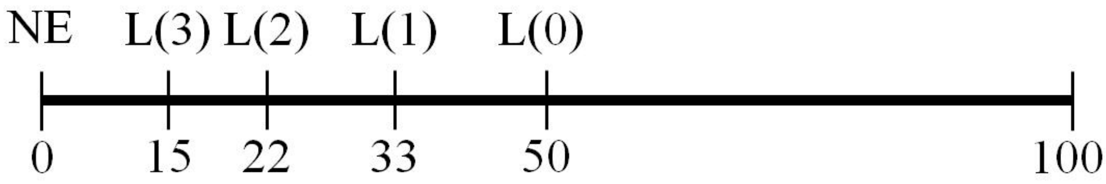


Fig. 4.1 | Steps of *k*-level reasoning in the p-Beauty contest game. A level 1 player $L(1)$ best responses to all naïve players that would randomly pick a number from the interval I [0,100] and chooses therefore 33. However, if everyone chooses 33, a level 2 player $L(2)$ would best respond to that by playing 22, and so on. A reasoning process with infinite steps converges to the unique (Nash) equilibrium *NE* of this game, in which everyone chooses 0.

Standard (economic) game theory suggests a unique (Nash) equilibrium in this game, in which – according to the fixed point argument – everyone chooses 0 as no one would deviate from this number given everyone else chooses 0 (see appendix D1 for details). This is similar to the notion of infinite steps of reasoning that converge to 0 as the applied steps k go to infinity ($= \frac{2}{3}^{k} \cdot 50$, with $k \rightarrow \infty$).

However, the question arises why empirical evidence shows that most people typically use quite low levels of reasoning, which usually varies between 1 and 3 steps (Coricelli & Nagel, 2009). Among other reasons, a rational guess (i.e. best response) in this game depends on one's own cognitive ability (Burnham, Cesarini, Johannesson, Lichtenstein, & Wallace, 2009; Carpenter, Graham, & Wolf, 2013; Devetag & Warglien, 2003), and additionally, on her beliefs about the average intelligence of her opponents (Agranov, Potamites, Schotter, & Tergiman, 2012; Camerer & Lovallo, 1999). That is, higher cognitive ability enables one to use more steps of reasoning, however, even if one is able to completely comprehend this game (i.e. infinite steps of reasoning), it is not a smart strategy to choose 0 when she believes that others might use only a few steps.

Therefore, in order to isolate the effect of oxytocin on strategic reasoning, it is necessary to control for these both factors, as oxytocin might influence general cognitive ability or one's beliefs about the others' sophistication. If ex ante neither general cognitive ability nor the assessment of others' intelligence is altered by oxytocin, it should be the case that the observed treatment effect can be traced back to a change in strategic sophistication. Hence, we controlled both factors by measuring subjects' general cognitive ability with a standard non-verbal assessment of their general intelligence, the Raven's advanced progressive matrices task (Raven, 1965), and additionally elicited their ex ante beliefs about their co-players performance in the *p*-Beauty contest game.

The hypothesis – oxytocin has a positive effect on strategic sophistication – was tested by employing a double-blind placebo controlled study design. Thus, this work compares the behavior of subjects in a competitive interactive game that received a single dose of intranasal self-administered oxytocin (N = 60) with that of subjects in a control group that received a placebo (N = 60). After 45 minutes of intake, participants were randomly assigned to groups of four and were asked to play the *p*-Beauty contest game for four rounds with feedback about the average and winning number after each round. The winner in each round receives a real monetary prize (i.e. 10 € per round), or a split of the prize in case of a tie (see *Methods*). Furthermore, all sessions were anonymous and fully computerized in order to ensure that participants could not sense or use any social cues of their opponents, since this work is intended to isolate the effect of oxytocin on purely strategic reasoning.

4.3 Results

4.3.1 Performance in the *p*-Beauty contest game

The results of the *p*-Beauty contest game show that guessed numbers were different between the two treatment conditions (oxytocin vs. placebo). In particular subjects in the oxytocin condition chose lower numbers in each round of the game (Fig. 4.2 A). Since four subjects interacted with each other in the same group and received feedback in each of the four rounds, I collapsed the data on group level by calculating means of each group for every round to obtain independent observations. In order to test whether participants under the influence of oxytocin chose lower numbers, I conducted a 2 (treatment: placebo/oxytocin) × 4 (rounds: 1, 2, 3 vs. 4) ANOVA with the first factor between groups. A significant effect of oxytocin on average guesses per group was found ($F(1,118) = 6.55$, $P = 0.011$). That is, subjects in the oxytocin group chose significantly lower numbers over all four rounds of the game. Replicating past results, subjects learned over time such that guesses in each group significantly declined over rounds ($F(3,116) = 55.24$, $P < 0.001$). However, an interaction variable of treatment × rounds was not significant ($F(3,116) = 0.08$, $P = 0.968$), showing that oxytocin has no round dependent or learning effect. In other words, guessed numbers in both treatment conditions declined almost parallel over the four rounds of the game, whereas guesses in the oxytocin group were significantly lower in each round of the game.

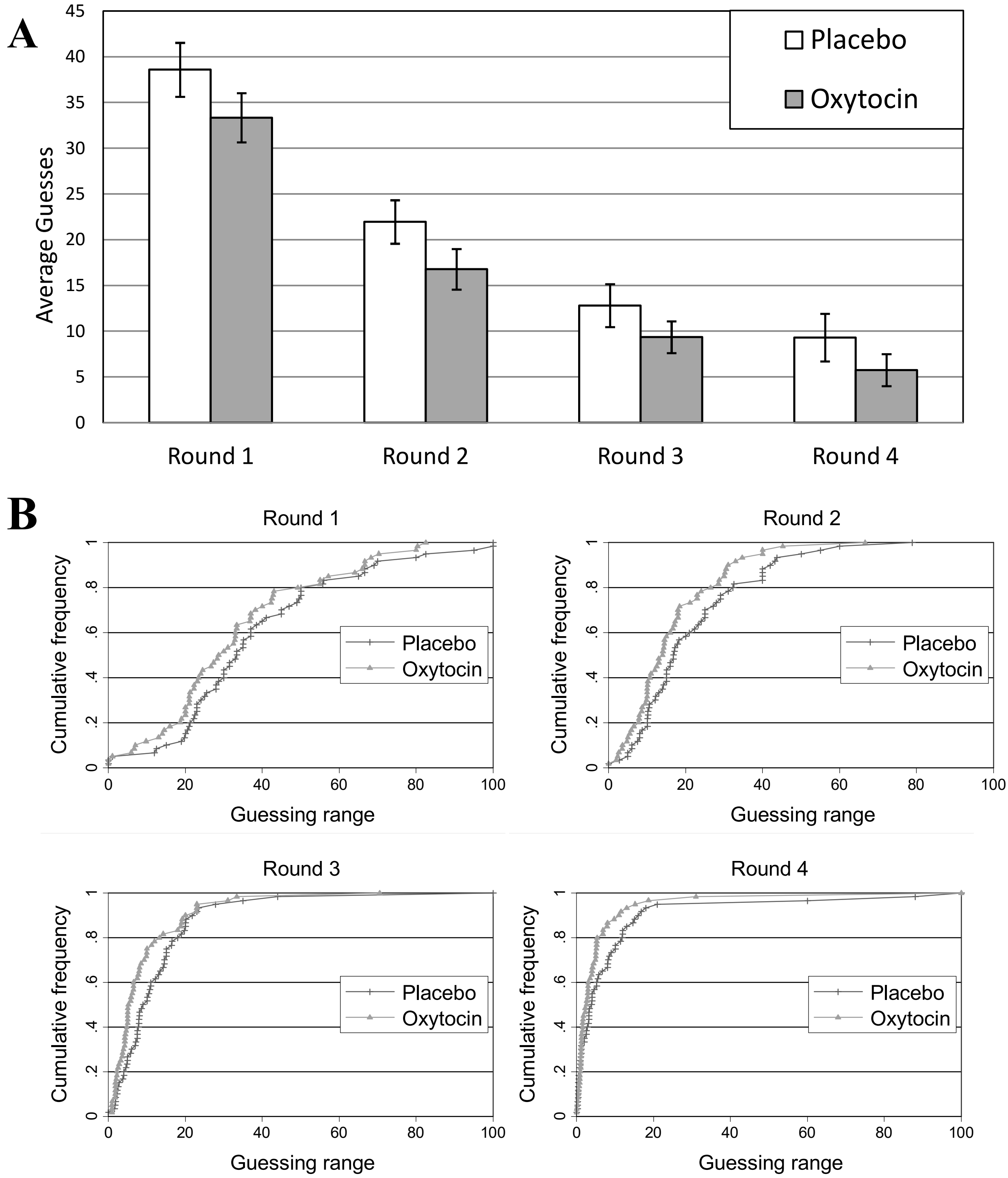


Fig. 4.2 | (*A*) Average guesses (± SE) in the *p*-Beauty contest game compared between oxytocin (N = 60) and placebo (N = 60). After the first round, data was collapsed on group level to obtain N = 15 independent observations per treatment condition. In both treatments guesses declined over time and converged to 0, the equilibrium of the game. (*B*) A cumulative distribution reveals a first order stochastic dominance of guesses in the placebo condition compared to oxytocin (N = 60 for each treatment). That is, guesses in the oxytocin treatment were lower for the whole range of the distribution in all four rounds.

A separate analysis of the decline rate with $d_t = guess_t \,/\, guess_{t-1}$ for each round t supports this notion. Except for the decline rate in d_2 with 0.57 (placebo) vs. 0.47 (oxytocin), the decline in all subsequent rounds was not significantly different compared between treatments (Mann-

Whitney U-test, $P = 0.093$ for $t = 2$, $P = 0.983$ for $t = 3$ and $P = 0.884$ for $t = 4$, two-sided; $N = 15$ in the oxytocin and $N = 15$ in the placebo treatment).

As stated before, in order to win the *p*-Beauty contest game participants would have to best respond to the average guesses of their co-players, which is equivalent to thinking one step more than others. Accordingly, a good measure of accurately predicting the behavior of others is characterized by the absolute deviation from the best possible response given the choices of all co-players. Thus, I calculated for each participant the absolute deviation of his guess from the winning number in his group to obtain a measure for properly estimating the sophistication of others. The absolute deviation a_t from the optimal response of subject i in period t is denoted by $a_t = |guess_{it} - ⅔\ mean_t|$, whereas $mean_t$ is the average of all guessed numbers in this group and $guess_{it}$ is the corresponding guess of subject i in this period. To test the effect of oxytocin on the accuracy of guessed numbers, I conducted a 2 (treatment: placebo/oxytocin) × 4 (rounds: 1, 2, 3 vs. 4) ANOVA with the first factor between groups. And indeed, participants in the oxytocin treatment were closer to the winning number in their group ($F(1,478) = 3.01$, $P = 0.0832$). Additionally, subjects learned over time such that the precision of guesses significantly rose over four rounds ($F(3,476) = 20.41$, $P < 0.001$). Again an interaction variable of treatment × rounds was not significant ($F(3,476) = 0.15$, $P = 0.929$), which is in line with the notion that oxytocin has no effect on learning in this game.

4.3.2 Applied thinking steps estimated with the Cognitive Hierarchy Model

The computation of the exact depths of reasoning in each treatment condition can be obtained by analyzing the first round of the experiment and by treating each guessed number per subject as an independent observation. In this round, participants began with lower entries in the oxytocin treatment, i.e. 33.3 on average, whereas the entries in the placebo treatment were on average 38.6 (Mann-Whitney U-test; $z = 1.404$, $P = 0.080$, one-sided, $N = 60$ in the oxytocin and $N = 60$ in the placebo group). To estimate the actual employed steps of thinking in the first round, I used the cognitive hierarchy model (Camerer et al., 2004) which constitutes a useful tool to describe behavior in this game (see appendix D2 for details). That is, I computed the actual depths of reasoning τ separately for each treatment condition. The estimation of the average reasoning level (see Table 4.1) revealed that participants under the influence of oxytocin applied 1.73 steps, whereas in the placebo treatment subjects reached a level of 0.90 (post-estimation Wald test; $z = 1.453$; $P = 0.073$, one-sided, $N = 60$ in the oxytocin and $N = 60$ in the placebo group). That is, the exogenous alteration of brain oxytocin levels resulted in a substantial enhancement of strategic reasoning, as the applied steps almost doubled.

	Average steps of iterations τ	
	Placebo group	*Oxytocin group*
Thinking steps in round one	***0.90***	***1.73***
	(0.38)	*(0.44)*

Table 4.1 | Average applied thinking steps τ compared between oxytocin and placebo. The average thinking steps τ of the first round were computed for each treatment condition (oxytocin vs. placebo). The (bootstrapped) standard errors are depicted in parentheses. The results show that iteration steps in the oxytocin group are almost twice as high as in the placebo group.

4.3.3 The interplay of oxytocin and cognitive ability in the *p*-Beauty contest game

The question arises whether this effect is due to a positive effect of oxytocin on general cognitive ability or rather due to its effect on strategic reasoning. To measure the impact of exogenous applied oxytocin on general cognitive ability, participants were asked to solve the Raven's advanced progressive matrices task (henceforth RPM), a standard non-verbal assessment of eductive ability (or general intelligence). In the analyses of this task, no statistical differences in average RPM scores were found between oxytocin and placebo (12.03 vs. 12; Mann-Whitney U-test; $z = -0.155$, $P = 0.876$, two-sided, $N = 60$ in the oxytocin and $N = 60$ in the placebo group), which indicates that the neuropeptide has no effect on subject's general intelligence. However, a regression analysis of guesses in the *p*-Beauty contest revealed that an increase in oxytocin levels, cognitive ability and the interaction between both variables are good predictors for average guesses per group (Table 4.2).

	Dependent Variable: Average guessed number per group	
	(1)	(2)
Treatment (Oxytocin=1 or Placebo=0)	−32.680**	−32.680**
	(12.907)	(12.963)
RPM Score	−1.384*	−1.384*
	(0.772)	(0.776)
Treatment x RPM Score	2.357**	2.357**
	(1.039)	(1.044)
Round		−9.356***
		(0.720)
Constant	37.277***	60.668***
	(10.043)	(10.487)
N	120	120
R^2	0.056	0.5868

Table 4.2 | Random effects regression of average guessed numbers. Mean of guesses per group is the dependent variable in random effects estimation. Negative coefficients indicate lower guesses and therefore more steps of iterations. Note that ***, **, * denote significance at the 1%, 5%, 10% level and standard errors clustered on group level are in parentheses.

Replicating past results, RPM scores and therefore cognitive ability are associated with lower guessed numbers, which indicates more steps of reasoning (see regression 1). In line with previous statistical tests, the treatment had a positive effect on strategic thinking, as subjects under the influence of oxytocin guessed significantly lower numbers. In regression 2, the control variable *Rounds* caused no changes in coefficients or their significance levels. It is interesting to note that the regression exhibits a positive interaction effect of the treatment with the RPM Score. This indicates that oxytocin had a positive effect on subjects with low cognitive ability, whereas this effect declined with increasing general intelligence.

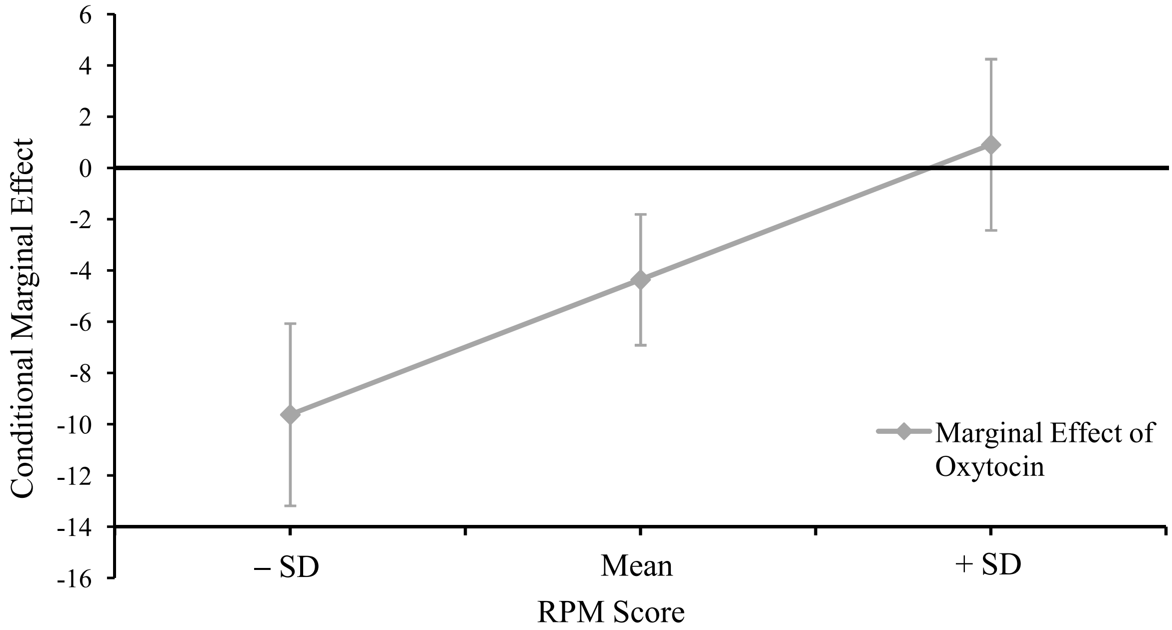


Fig. 4.3 | The effect of oxytocin conditional on general cognitive ability measured with the RPM (displayed ± SE). Subjects who exhibit relatively low (– SD) or average (Mean) general intelligence show a significant improvement of their strategic reasoning abilities under the influence of oxytocin. However, the effect is insignificant for subjects with relatively high general intelligence.

Fig. 4.3 illustrates this pattern in which the conditional marginal effect of oxytocin is depicted for three levels of general intelligence, i.e. mean and one standard deviation below or above average RPM score. Accordingly, the influence of oxytocin is statistically insignificant on subjects with one SD above average RPM score (z = 0.27, P = 0.785, two-sided). However, the effect was greatest for subjects who achieved low RPM scores (z = –2.70, P = 0.007, two-sided) and moderate for subjects with average RPM scores (z = –1.71, P = 0.088, two-sided).

4.3.4 Beliefs about winning chances

Forming correct beliefs about the average strategic sophistication of all opponents (i.e. their average reasoning level) is essential to win this game. As it has been shown by previous studies, subjects who believe to play against more sophisticated players choose significantly lower numbers. In order to rule out the possibility that oxytocin might wrongly bias beliefs about the strategic sophistication of others, participants' beliefs regarding their performance in relation to their co-players were elicited. Since the p-Beauty contest is a fixed sum game in which (at least) one group-member has to win in every round, it is apparent that a lower self-assessment is equivalent to a higher evaluation of others' abilities. Thus, winning exactly once, for instance, is the same as saying the others will win 3 times (splitting the price can be ignored due to the very rare case of a tie with 3 out of 120 in the experiment). Accordingly, before the beginning of the first task participants were asked to estimate their winning frequency (ranging from 0 to 4 times). If their guess was right, participants would receive an extra bonus of 1.60 Euros. However, I found no statistical difference in beliefs regarding predicted winning frequencies between treatment conditions. In the oxytocin treatment, subjects predicted an average winning frequency of 1.51 compared to 1.41 in the placebo treatment (Mann-Whitney U-test; $z = 0.649$, $P = 0.516$; two-sided test). Thus, no support was found for the hypothesis that oxytocin influences overconfidence and accordingly, the assessment of others.

4.3.5 Response times in the *p*-Beauty contest game

Provided that strategic reasoning, like reasoning in general, is a slow and effortful process, it is straightforward to presume that response times in the *p*-Beauty contest game should differ between treatment conditions. That is, as subjects in the oxytocin condition exhibit more steps of reasoning, one would expect that those subjects think harder and therefore longer in this strategic task. And indeed, the average response time (collapsed on group level) amounts to 51.1 seconds in the oxytocin and to 35.2 seconds in the placebo treatment (Mann-Whitney U-test; $z = -2.219$, $P = 0.026$, two-sided, $N = 15$ in the oxytocin and $N = 15$ in the placebo group). That is, participants under the influence of oxytocin were willing to take about 45% more time to reason about their next number, indicating a longer and therefore deeper process of thinking (Fig. 4.4). However, it is necessary to compare response times with a neutral task in order to rule out a general extension of response times due to oxytocin. Therefore, I used a neutral task with no relation to processes of strategic sophistication or *k*-level reasoning. The response times of the belief questionnaire seem to be an ideal neutral comparison, since this task was also incentivized and contained no elements of strategic interaction.

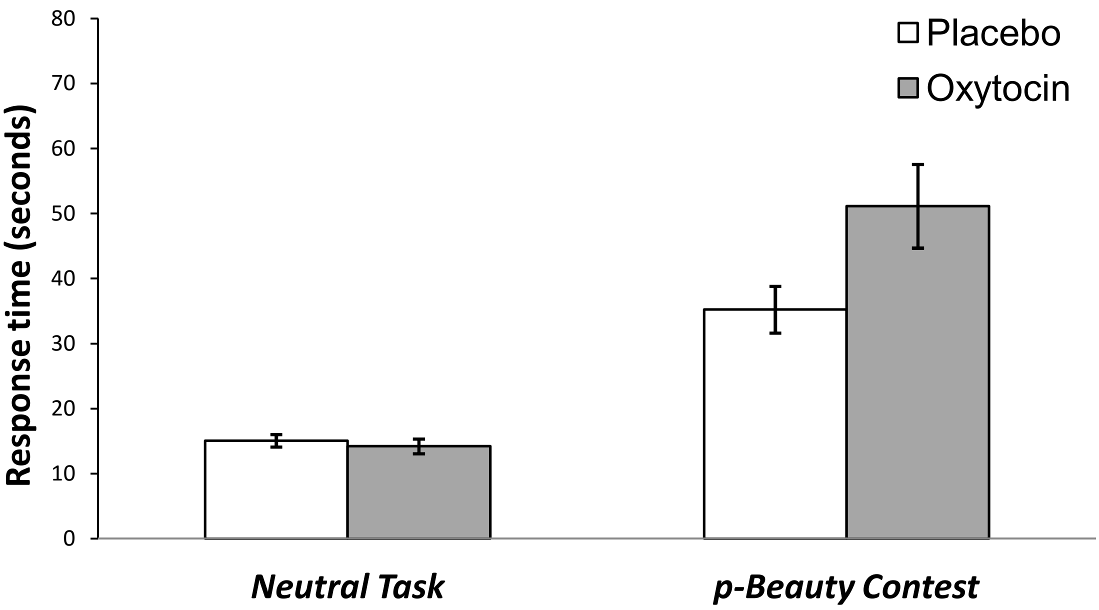


Fig. 4.4 | Analysis of response times per decision as a function of treatment (oxytocin vs. placebo) and task (Neutral vs. *p*-Beauty contest). Average response time (± SE) (in seconds) when choosing a number compared between oxytocin (N = 15) and placebo (N = 15) treatment differed, whereas the average response time (± SE) per decision in the neutral task (belief questionnaire) was statistically identical between oxytocin (N = 60) and placebo (N = 60).

And indeed, I found no difference in response times between treatment conditions for this neutral questionnaire (mean = 15.1 in placebo vs. 14.2 in oxytocin group; Mann-Whitney U-test; z = 1.376, P = 0.169, two-sided, N = 60 in the oxytocin and N = 60 in the placebo group). Consequently, longer response times were caused by more steps of reasoning in the oxytocin treatment, which in turn led to more sophisticated guesses.

4.4 Discussion

This work is able to show that oxytocin has a positive effect on strategic reasoning which is most pronounced for subjects with lower general cognitive ability. These findings contribute to previous studies that demonstrated the important role of oxytocinergic systems in cognitive empathy, the ability to put oneself emotionally in the shoes of another person (Domes et al., 2007; Rodrigues et al., 2009; Schulze et al., 2011; Shamay-Tsoory et al., 2013). Thus, this work extends this notion by providing evidence for a positive effect of oxytocin on strategic sophistication in purely rational interactions. In particular, I used the cognitive hierarchy model to provide a computational measure of levels of strategic thinking (Camerer et al., 2004), and according to the results I found that on average oxytocin almost doubled the utilized level of reasoning.

The analysis regarding the impact of general cognitive ability on strategic reasoning is in line with previous studies that found a strong relationship between general intelligence and success in the *p*-Beauty contest game (Burnham et al., 2009; Carpenter et al., 2013). The findings demonstrate a similar association of higher general intelligence with lower and therefore more sophisticated guessed numbers. However, I found no evidence that oxytocin has a positive effect on general or eductive intelligence, which was measured with the Raven's advanced progressive matrices task. Furthermore, I found a remarkable interaction between oxytocin and general intelligence in the experiment. Participants with higher cognitive ability seemed to hardly respond to the application of oxytocin, as they showed no difference in behavior compared to the placebo condition. Conversely, oxytocin significantly improved the strategic sophistication of participants with average or low general intelligence. The conditional effect of oxytocin on strategic sophistication may be due to the fact that intelligent subjects already exhibited quite low guesses, so that their sophistication could hardly be improved. Thus, guesses of subjects with low cognitive ability showed a higher potential of improvement.

Another important aspect of mentalizing is the ability to form correct beliefs about the sophistication of others. Especially in strategic interactions, it is indispensable to predict the cognitive abilities of co-players, as beliefs about the opponents' level of sophistication dramatically changes optimal behavior (Agranov et al., 2012). Therefore, in this experiment any provision of social information, which could be considered as a sign of the average level of sophistication, was carefully avoided. This is of particular importance, since previous findings indicate that participants under the influence of oxytocin react stronger to social information (Bartz et al., 2011; Declerck et al., 2010). To obtain an additional control, I elicited the beliefs of participants regarding their co-players' performance, and tried to capture any possible shift in beliefs about the average sophistication of others. The results provide no evidence that oxytocin shaped the assessment of the average sophistication level. I therefore reason that this can be ruled out as an alternative explanation of the observed treatment effect on strategic thinking.

Accordingly, as the two driving factors that explain behavior in the *p*-Beauty contest game, i.e. cognitive ability and belief formation, were not influenced by oxytocin, I conclude that the observed treatment effect was caused by an improvement of strategic thinking.

However, if the neuropeptide improves *k*-level thinking, it is straightforward to presume that subjects in the oxytocin condition should show longer response times, since strategic reasoning, like reasoning in general, is a slow and effortful process rather than a fast and impulsive one (Frederick, 2005; Metcalfe & Mischel, 1999). This view seems to be supported by evidence showing that brain areas related to higher cognitive functions are involved in strategic reasoning and mentalizing (Amodio & Frith, 2006; Carrington & Bailey, 2009; Coricelli & Nagel, 2009), and longer decision times are associated with greater performance in the *p*-Beauty contest game (Chong, Camerer, & Ho, 2005; Kocher & Sutter, 2006; Rubinstein, 2007). And indeed, participants in the oxytocin condition took 45% more time to decide on their next guessed number, whereas response times in a neutral task showed no difference. As response times were only different in a task that involved *k*-level thinking, I conclude that more steps of reasoning in the oxytocin treatment induced longer decision times.

4.5 Conclusion

As it has been shown before, oxytocin plays a key role in manifold elements of social interactions and especially in the maintenance of social groups. For instance, it modulates many basic elements that characterize social groups like trust (Kosfeld et al., 2005), cognitive empathy (Domes et al., 2007; Rodrigues et al., 2009), in-group favoritism (De Dreu et al., 2011) and parochial altruism (De Dreu et al., 2010). As social beings, humans sustain societies, for which it is indispensable to infer the strategies of others which paves the way for coordination, cooperation (De Dreu et al., 2010) and trust (Kosfeld et al., 2005; Zak et al., 2005) among group members. In view of that, it is not surprising that a dysfunction of the oxytocinergic system is related to several disorders characterized by impairments in social interaction and mentalizing abilities, including particular forms of autism spectrum disorder (Andari et al., 2010; Guastella et al., 2010; Jacob et al., 2007; Modahl et al., 1998) and schizophrenia (Goldman, Marlow-O'Connor, Torres, & Carter, 2008; Pedersen et al., 2011). In light of this evidence, the oxytocinergic system may constitute a necessity to enable the evolvement of larger societies. This would be in line with the *social brain hypothesis* (Dunbar, 1998; Dunbar & Shultz, 2007) as larger societies with a tremendous set of social interactions require more cognitive capacity with specialized neural systems. Therefore, oxytocin might have played a key role in the evolvement of mentalizing abilities of anthropoid primates (Dunbar & Shultz, 2007; Seyfarth & Cheney, 2013) who needed to process an increasing number of social information in even more complex forms of social interactions. This work contributes to this notion by showing that strategic reasoning (or strategic mentalizing) is partially encoded in special neural substrates related to the oxytocinergic system (Amodio & Frith, 2006; Coricelli & Nagel, 2009).

4.6 Materials and Methods

4.6.1 Participants

A total of 120 male subjects from different universities in Munich were recruited via ORSEE (Greiner, 2004) and participated in seven experimental sessions at the Munich Experimental Laboratory (MELESSA). Participants were between 18 and 35 years old (mean age ± s.d., 23.1 ± 3.3 years). Exclusion criteria for participation were significant medical or psychiatric disorder, medication, smoking more than 15 cigarettes per day, and drug or alcohol abuse. Subjects were instructed to abstain from alcohol, smoking and caffeine for 24 h, and from food for 1 h before the experiment. At the time of recruitment, participants were informed that the experiment is about the influence of oxytocin on economic decision making. The study was approved by the ethics committee of the medical department at the University of Munich. All participants provided written, informed consent before participation.

4.6.2 Substance administration

I used a double-blind, placebo-controlled study design to compare behavior in a group of subjects that received a single intranasal dose of 24 IU oxytocin (Syntocinon-Spray, Sigma-Tau; 3 puffs per nostril, each with 4 IU of oxytocin) with that of subjects in a control group that received a placebo. Subjects were randomly assigned to the two groups and received oxytocin or a placebo 45 minutes before the relevant part of the experiment. In one experimental session either oxytocin or placebo was administered. The placebo contained all inactive ingredients except for the neuropeptide (Born et al., 2002; Gossen et al., 2012).

4.6.3 Behavioral experiment and questionnaires

Subjects received neutrally written instructions, which were read aloud (see appendix D3 for details). They knew that the experiment consisted of several parts which were independent of each other. Instructions for one part only followed after the previous part was accomplished. Decisions were anonymously made on PC screens in the experimental laboratory. 45 minutes after substance administration, the relevant part of the experiment started, in which subjects were asked to conduct an incentivized version of the Raven's progressive matrices task (RPM) with five minutes time to solve as many questions as possible (Raven, 1965). The RPM comprises 36 items of diagrammatic puzzles which were presented in progressive difficulty. In each puzzle, one piece is missing and subjects have to find the correct piece to complete the diagrammatic pattern. This makes the RPM an ideal tool to measure subject's (non-strategic) reasoning ability in a complex non-verbal task. For every correctly solved question participants received 0.5 MU, however, in order to prevent random guessing, the same amount was deduced for every incorrect answer. After this, all subjects were randomly assigned to groups of four and were asked to play the *p*-Beauty contest game. Particularly, subjects had to choose a real number from the interval $I \equiv [0, 100]$. The one, whose number was closest to 2/3 times the mean of all chosen numbers, won 30 MUs or a split of the prize in case of a tie. However, just before the actual task began, participants' beliefs about the performance of their group members were elicited. For the correct prediction, subjects received an extra bonus of 5 MU. Then, subjects played the *p*-Beauty contest game for 4 rounds with the same composition of the group over the course of the game. The use of a calculator was allowed. Feedback consisted of the mean of all numbers in the respective group, the target number and whether the subject had won or not.

In the end of the experiment, all participants received a show up fee of 4 € and their payoff corresponding to their success in the experiment. Each MU earned in the experiment was worth 0.33 €. The whole experiment took about 90 minutes and was programmed in z-Tree, an experimental open source software (Fischbacher, 2007).

Appendix D

D1. Standard game theoretic predictions of the *p*-Beauty contest game

In the *p*-Beauty contest game, standard game theory suggests (for $p = 2/3$) that everyone chooses 0, the unique (Nash) equilibrium of this game (Nagel, 1995).

Proof by the iterative elimination of (weakly) dominated strategies:

If all players are rational and it is common knowledge that all players are rational, then a rational player will exclude any number greater than $p \cdot 100$ from his choice set, because all numbers of the interval $[p \cdot 100, 100]$ are (weakly) dominated by $p \cdot 100$. However, if everyone knows that everyone excludes $[p \cdot 100, 100]$, then players will exclude any number greater than $p^2 \cdot 100$, and so on. If players apply infinite steps of iterated elimination of weakly dominated strategies, all numbers greater than $p^k \cdot 100$, with $k \to \infty$, will be excluded from their choice set. The only remaining strategy is 0, the unique equilibrium of this game.

D2. The Assumptions of the Cognitive Hierarchy Model

The cognitive hierarchy model (Camerer et al., 2004) assumes that each player is a step-k player, and has accurate beliefs about the proportions of players doing h steps of thinking. This proportion of players doing h steps is given by:

$$g_k(h) = f(h) / \sum_{l=0}^{k-1} f(l), \forall h < k$$

Note that everyone assumes that there are no players applying higher steps of reasoning as they do, such that $g_k(h) = 0$, $\forall\ h \geq k$. Furthermore, step-0 players are not thinking strategically and, for simplicity, equally randomize over all strategies.

Then, a step-k player best responds to all other players according to his beliefs about the proportions of all lower types, given by $f(0), f(1), f(2), \ldots f(k-1)$.

To obtain plausible properties of the distribution $f(k)$, it is necessary to define further requirements of it. For instance, the distribution of $f(k)$ should be discrete, since the thinking steps k are integers. And of course, the relative proportion $f(k) / f(k-1)$ should decline with k, as more steps of thinking need more cognitive capacity, which is by assumption a limited resource. If one assumes that the decline of $f(k) / f(k-1)$ is proportional to $1/k$, then $f(k)$ is the Poisson distribution, and is given by:

$$f(k) = \frac{e^{-\tau}\tau^k}{k!}$$

Note that τ represents the mean of employed thinking steps and its variance. This gives the possibility to estimate the average steps of thinking τ by using GMM estimation techniques.

GMM Estimation of the reasoning parameter τ

As step-0 players randomize over the whole interval of [0,100], step-1 players will best-respond to them by choosing 33 (for $p = ⅔$). However, step-2 players will best-respond to a linear combination of step-1 and step-0 players. That is, step-2 players will weight 33 and 50 by the frequencies of step-1 and step-0 players in the population. Thus, step-2 players will best respond to $(50 \cdot f(0) + 33 \cdot f(1)) / (f(0) + f(1))$. As $f(k)$ is Poisson distributed, a step-2 player will best respond to $(50 + 33 \cdot \tau) / (1 + \tau)$.

Consequently, I can estimate τ separately for each treatment condition (oxytocin and placebo) by using the generalized method of moments. That is, for each treatment group I minimize the difference between the predicted mean and the actual mean of guesses.

D3. Experimental Instructions (Translated from German)

Welcome to the experiment and thank you for your participation! From now on, please do not talk to other participants and please switch off your phone!

General Procedure

The purpose of this experiment is to study decision making. You can earn money during this experiment which will be paid out to you in cash at the end of the experiment. During the experiment you and the other participants will make decisions. Your own decisions as well as those of the other participants will affect your payment according to the rules explained in the following. The duration of the experiment is 120 minutes. If you should have any questions, please raise your hand. One of the experimenters will come to you and answer your questions in private. During the experiment we do not speak of Euros, but of Guilders. Thus, your earnings during the experiment will be calculated in Guilders. At the end of the experiment your earnings in Guilders will be converted into Euros according to the following exchange rate: 3 Guilders = 1 Euro and 1 Guilder = 33 Eurocent.

For your punctual arrival you additionally receive 12 Guilders. At the end of the experiment you will be paid your total earnings from the experiment in private and in cash. While you make your decisions a clock will count down at the upper right corner of the computer screen. This countdown serves as an orientation for how much time you should need to take the decision. However, usually you can take more time to make your decision. If it is not possible to take more time than indicated, we will inform you in advance.

Anonymity

Your decisions over the course of the experiment are anonymous. You will not receive personal information about the other participants of the experiment. All data elicited will be encoded and will be treated confidentially. At the end of the experiment you will sign a confirmation which says that you have received the payment. This confirmation only serves accounting purposes for our sponsor. Our sponsor will not receive any individual data from the experiment either.

Utilities

For this experiment you are given a pen and a scratch paper. We kindly ask you to leave them at your place after the experiment.

The experiment

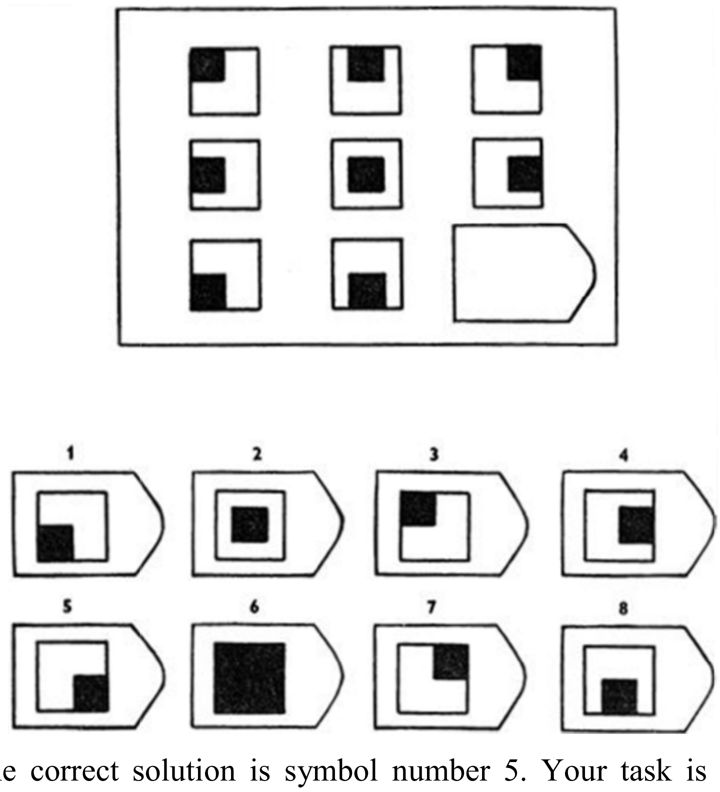
For each part you will receive separate instructions in written and/or verbal form. The individual parts are independent from each other, this means decisions made in one part of the experiment will not affect later parts. Your total profit will be the sum of the individual profits and the 12 Guilders you receive for punctual arrival. At the end of the experiment you will be informed about your total profit and paid individually in cash.

Application of nasal sprays

In the first part we ask you to apply yourself a nasal spray. Neither you nor the attending experimenters know whether you are given oxytocin or a non-effective aqueous solution (placebo). The sprays are distributed by a medical professional and the experimenter. Please refrain from application until asked to do so. All participants should apply the spray at the same time. Please do not communicate with your neighbours. If you have any private questions, please raise your hand. An experimenter will answer your questions privately.

Raven advanced progressive matrices task (RPM)

In the following, all participants are requested to solve an assignment, which will be explained more precisely in the following: On your screen a matrix (i.e. a rectangular arrangement of different symbols) will appear in a framed box. The matrix has 3 columns and 3 rows. The symbol in the lower right corner is missing. Beneath the matrix, there are 8 symbols; exactly one of them fits schematically in the lower right corner of the matrix. An example is given below:



In this example the correct solution is symbol number 5. Your task is to find the correct symbol. After you made your choice, a new matrix including 8 new symbols will appear. Again, exactly one of the 8 symbols fits in the lower right corner of the matrix. You have to

choose a symbol for each matrix. You cannot move on to the next matrix without making a choice. Moreover you will not be able to go back to the previous matrix and change your choice after you have confirmed it. You have 5 minutes (300 seconds) to solve as many matrixes as possible. Your remaining time will be displayed in the upper right corner of your screen. On the left bottom of the screen, you will see how many matrices you have solved correctly and incorrectly so far, and whether your last choice was correct or wrong. The difficulty of the matrices will gradually increase over time. Your payment in this part is calculated as follows: You will receive 0.5 Guilders for each matrix you have solved correctly. For each matrix you solve wrongly, 0.5 Guilders will be deducted from your earnings. Thus, your total income is denoted by: 0.5 x (number of correctly solved matrices – number of incorrectly solved matrices). If you have solved more matrices incorrectly than correctly, you will receive 0 Guilders in this part. That is, a negative income is not possible.
Before starting with the actual task, you will be able to solve some trial matrices, in order to familiarize you with this kind of task. The trial matrices will not be paid.

The *p*-Beauty contest game *(distributed after the completion of previous part)*
First, all participants will be randomly allocated to groups of 4. Then, each person chooses a number from the interval of [0,100] by typing the number in the corresponding box on the screen. It is also possible to choose 0 or 100. The number wins which is closest to two-thirds times the mean of all chosen numbers in your group. (Target number = 2/3 x mean of all chosen numbers in your group). In mathematical terms, the target number is denoted by:

$$\bar{x} = \frac{2}{3}\frac{\sum_{i=1}^{n} x_i}{n}$$

That is, each person in your group types (privately) a number in the computer, whereas the person in your group wins whose number is closest to the target number. If your chosen number is closest to the target number (in your group), you win. This task will be repeated for 4 periods, whereas the composition of your group remains constant over all 4 periods. The winner in the respective period receives 30 Guilders. In case of a tie, the prize will be divided equally among all winners. After each period you will receive feedback about the target number and whether you have won or not. Please note that 2 decimal places are allowed.
Using the calculator: If you want to make some calculations, just click on the calculator icon on the bottom of your screen which opens the Windows-calculator. (Please keep in mind the “order of operation rule”).

Bibliography

Abeler, J., Becker, A., & Falk, A. (2014). Representative evidence on lying costs. *Journal of Public Economics*, *113*(0), 96–104.

Agranov, M., Potamites, E., Schotter, A., & Tergiman, C. (2012). Beliefs and endogenous cognitive levels: An experimental study. *Games and Economic Behavior*, *75*(2), 449–463.

Akerlof, G. A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The quarterly journal of economics*, *84*(3), 488–500.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*(4), 268–277.

Andari, E., Duhamel, J.-R., Zalla, T., Herbrecht, E., Leboyer, M., & Sirigu, A. (2010). Promoting social behavior with oxytocin in high-functioning autism spectrum disorders. *Proceedings of the National Academy of Sciences*, *107*(9), 4389–4394.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182.

Bartz, J., Simeon, D., Hamilton, H., Kim, S., Crystal, S., Braun, A., . . . Hollander, E. (2011). Oxytocin can hinder trust and cooperation in borderline personality disorder. *Social Cognitive and Affective Neuroscience*, *6*(5), 556–563.

Bartz, 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.

Baumgartner, 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.

Belot, M., & Schröder, M. (2013). Sloppy work, lies and theft: A novel experimental design to study counterproductive behaviour. *Journal of Economic Behavior & Organization*, *93*(0), 233–238.

Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature*, *442*(7105), 912–915.

Bethlehem, R. A. I., Baron-Cohen, S., van Honk, J., Auyeung, B., & Bos, P. A. (2014). The oxytocin paradox. *Frontiers in Behavioral Neuroscience*, *8*, 48.

Bhatt, M., & Camerer, C. F. (2005). Self-referential thinking and equilibrium as states of mind in games: fMRI evidence. *Games and Economic Behavior*, *52*(2), 424–459.

Born, J., Lange, T., Kern, W., McGregor, G. P., Bickel, U., & Fehm, H. L. (2002). Sniffing neuropeptides: a transnasal approach to the human brain. *Nature neuroscience*, *5*(6), 514–516.

Burnham, T. C., Cesarini, D., Johannesson, M., Lichtenstein, P., & Wallace, B. (2009). Higher cognitive ability is associated with lower entries in a p-beauty contest. *Journal of Economic Behavior & Organization*, *72*(1), 171–175.

Camerer, C., Loewenstein, G., & Prelec, D. (2005). Neuroeconomics: How Neuroscience Can Inform Economics. *Journal of Economic Literature*, *43*(1), 9–64.

Camerer, C., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *The American economic review*, *89*(1), 306–318.

Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2004). A Cognitive Hierarchy Model of Games. *The quarterly journal of economics*, *119*(3), 861–898.

Carpenter, J., Graham, M., & Wolf, J. (2013). Cognitive ability and strategic sophistication. *Games and Economic Behavior*, *80*, 115–130.

Carrington, S. J., & Bailey, A. J. (2009). Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Human brain mapping*, *30*(8), 2313–2335.

Charness, G., & Sutter, M. (2012). Groups Make Better Self-Interested Decisions. *The Journal of Economic Perspectives*, *26*(3), 157–176.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics*, *14*(1), 47–83.

Chong, J.-K., Camerer, C., & Ho, T.-H. (2005). Cognitive Hierarchy: A Limited Thinking Theory in Games. In R. Zwick & A. Rapoport (Eds.), *Experimental Business Research* (pp. 203–228). Springer US.

Churchland, P. S., & Winkielman, P. (2012). Modulating social behavior with oxytocin: How does it work? What does it mean? *Oxytocin, Vasopressin and Social Behavior*, *61*(3), 392–399.

Cialdini, R. B., & Goldstein, N. J. (2004). Social Influence: Compliance and Conformity. *Annual Review of Psychology*, *55*(1), 591–621.

Cohn, A., Fehr, E., & Marechal, M. A. (2014). Business culture and dishonesty in the banking industry. *Nature*, *516*(7529), 86–89.

Cooper, D. J., & Kagel, J. H. (2005). Are Two Heads Better than One? Team versus Individual Play in Signaling Games. *The American economic review*, *95*(3), 477–509.

Coricelli, G., & Nagel, R. (2009). Neural correlates of depth of strategic reasoning in medial prefrontal cortex. *Proceedings of the National Academy of Sciences*, *106*(23), 9163–9168.

De Dreu, C. K. W., Greer, L. L., Handgraaf, M. J. J., Shalvi, S., Van Kleef, G. A., Baas, M., . . . Feith, S. W. W. (2010). The Neuropeptide Oxytocin Regulates Parochial Altruism in Intergroup Conflict Among Humans. *Science*, *328*(5984), 1408–1411.

De Dreu, C. K. W., Greer, L. L., Van Kleef, G. A., Shalvi, S., & Handgraaf, M. J. J. (2011). Oxytocin promotes human ethnocentrism. *Proceedings of the National Academy of Sciences*, *108*(4), 1262–1266.

De Quervain, D. J.-F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., & Fehr, E. (2004). The neural basis of altruistic punishment. *Science (New York, N.Y.)*, *305*(5688), 1254–1258.

Declerck, C., Boone, C., & Kiyonari, T. (2013). The effect of oxytocin on cooperation in a prisoner's dilemma depends on the social context and a person's social value orientation. *Social Cognitive and Affective Neuroscience*, Declerck, C. H., Boone, C., & Kiyonari, T. (2010). Oxytocin and cooperation under conditions of uncertainty: the modulating role of incentives and social information. *Hormones and Behavior*, *57*(3), 368–374.

Devetag, G., & Warglien, M. (2003). Games and phone numbers: Do short-term memory bounds affect strategic behavior? *Journal of Economic Psychology*, *24*(2), 189–202.

Ditzen, B., Schaer, M., Gabriel, B., Bodenmann, G., Ehlert, U., & Heinrichs, M. (2009). Intranasal Oxytocin Increases Positive Communication and Reduces Cortisol Levels During Couple Conflict. *Biological Psychiatry*, *65*(9), 728–731.

Dolen, G., Darvishzadeh, A., Huang, K. W., & Malenka, R. C. (2013). Social reward requires coordinated activity of nucleus accumbens oxytocin and serotonin. *Nature*, *501*(7466), 179–184.

Domes, G., Heinrichs, M., Michel, A., Berger, C., & Herpertz, S. C. (2007). Oxytocin improves "mind-reading" in humans. *Biological Psychiatry*, *61*(6), 731–733.

Donaldson, Z. R., & Young, L. J. (2008a). Oxytocin, vasopressin, and the neurogenetics of sociality. *Science*, *322*(5903), 900–904.

Donaldson, Z. R., & Young, L. J. (2008b). Oxytocin, Vasopressin, and the Neurogenetics of Sociality. *Science*, *322*(5903), 900–904.

Dunbar, R. (1998). The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews*, *6*(5), 178–190.

Dunbar, R., & Shultz, S. (2007). Evolution in the social brain. *Science*, *317*(5843), 1344–1347.

Egas, M., & Riedl, A. (2008). The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1637), 871–878.

Eisenegger, C., Haushofer, J., & Fehr, E. (2011). The role of testosterone in social interaction. *Trends in Cognitive Sciences*, *15*(6), 263–271.

Falk, A., Fehr, E., & Fischbacher, U. (2005). Driving Forces Behind Informal Sanctions. *Econometrica*, *73*(6), 2017–2030.

Falk, A., & Szech, N. (2013). Morals and Markets. *Science*, *340*(6133), 707–711.

Faravelli, M., Friesen, L., & Gangadharan, L. (2015). Selection, tournaments, and dishonesty. *Journal of Economic Behavior & Organization*, *110*(0), 160–175.

Fehr, E., & Fischbacher, U. (2004). Social norms and human cooperation. *Trends in cognitive sciences*, *8*(4), 185–190.

Fehr, E., & Gächter, S. (2000). Cooperation and Punishment in Public Goods Experiments. *American Economic Review*, *90*(4), 980–994.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137–140.

Fehr, E., & Schmidt, K. M. (1999). A Theory of Fairness, Competition, and Cooperation. *The quarterly journal of economics*, *114*(3), 817–868.

Feri, F., Irlenbusch, B., & Sutter, M. (2010). Efficiency Gains from Team-Based Coordination—Large-Scale Experimental Evidence. *The American economic review*, *100*(4), 1892–1912.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–178.

Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies In Disguise—An Experimental Study On Cheating. *Journal of the European Economic Association*, *11*(3), 525–547.

Frederick, S. (2005). Cognitive Reflection and Decision Making. *The Journal of Economic Perspectives*, *19*(4), 25–42.

Frey, B. S. (1994). Direct Democracy: Politico-Economic Lessons from Swiss Experience. *The American economic review*, *84*(2), 338–342.

Gino, F., Ayal, S., & Ariely, D. (2009). Contagion and Differentiation in Unethical Behavior: The Effect of One Bad Apple on the Barrel. *Psychological Science*, *20*(3), 393–398.

Goldman, M., Marlow-O'Connor, M., Torres, I., & Carter, C. S. (2008). Diminished plasma oxytocin in schizophrenic patients with neuroendocrine dysfunction and emotional deficits. *Schizophrenia research*, *98*(1), 247–255.

Gossen, 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.

Greiner, B. (2004). An online recruitment system for economic experiments.

Grillon, C., Krimsky, M., Charney, Vytal, K., Ernst, M., & Cornwell, B. (2012). Oxytocin increases anxiety to unpredictable threat. *Molecular psychiatry*, *18*(9), 958–960.

Groppe, S. E., Gossen, A., Rademacher, L., Hahn, A., Westphal, L., Gründer, G., & Spreckelmeyer, K. N. (2013). Oxytocin Influences Processing of Socially Relevant Cues in the Ventral Tegmental Area of the Human Brain. *Oxytocin and Autism*, *74*(3), 172–179.

Gruter, M., & Masters, R. D. (1986). Ostracism as a social and biological phenomenon: An introduction. *Ethology and Sociobiology*, *7*(3–4), 149–158.

Guastella, A. J., Einfeld, S. L., Gray, K. M., Rinehart, N. J., Tonge, B. J., Lambert, T. J., & Hickie, I. B. (2010). Intranasal oxytocin improves emotion recognition for youth with autism spectrum disorders. *Biological Psychiatry*, *67*(7), 692–694.

Hamilton, B. H., Nickerson, J. A., & Owan, H. (2003). Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation. *Journal of Political Economy*, *111*(3), p 465-497.

Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proceedings of the National Academy of Sciences*, *105*(18), 6741–6746.

Hicks, R., & Tingley, D. (2011). Causal mediation analysis. *Stata Journal*, *11*(4), 605–619.

Hopfensitz, A., & Reuben, E. (2009). The Importance of Emotions for the Effectiveness of Social Punishment. *The Economic Journal*, *119*(540), 1534–1559.

Houser, D., Vetter, S., & Winter, J. (2012). Fairness and cheating. *European Economic Review*, *56*(8), 1645–1655.

Jacob, S., Brune, C. W., Carter, C. S., Leventhal, B. L., Lord, C., & Cook, E. H. (2007). Association of the oxytocin receptor gene (OXTR) in Caucasian children and adolescents with autism. *Neuroscience Letters*, *417*(1), 6–9.

Kerr, N. L., & Tindale, R. S. (2004). Group Performance and Decision Making. *Annual Review of Psychology*, *55*(1), 623–655.

Keynes, J. M. (2006). *General theory of employment, interest and money*: Atlantic Publishers & Dist.

Knack, S., & Keefer, P. (1997). Does Social Capital Have an Economic Payoff? A Cross-Country Investigation. *The quarterly journal of economics*, *112*(4), 1251–1288.

Kocher, M. G., Strauß, S., & Sutter, M. (2006). Individual or team decision-making—Causes and consequences of self-selection. *Games and Economic Behavior*, *56*(2), 259–270.

Kocher, M. G., & Sutter, M. (2005). The Decision Maker Matters: Individual Versus Group Behaviour in Experimental Beauty-Contest Games *The Economic Journal*, *115*(500), 200–223.

Kocher, M. G., & Sutter, M. (2006). Time is money—Time pressure, incentives, and the quality of decision-making. *Journal of Economic Behavior & Organization*, *61*(3), 375–392.

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, *435*(7042), 673–676.

Kugler, T., Kausel, E. E., & Kocher, M. G. (2012). Are groups more rational than individuals? A review of interactive decision making in groups. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(4), 471–482.

Laughlin, P. R., & Ellis, A. L. (1986). Demonstrability and social combination processes on mathematical intellective tasks. *Journal of Experimental Social Psychology*, *22*(3), 177–189.

Laughlin, P. R., Hatch, E. C., Silver, J. S., & Boh, L. (2006). Groups perform better than the best individuals on letters-to-numbers problems: Effects of group size. *Journal of Personality and Social Psychology*, *90*(4), 644–651.

MacDonald, K., & MacDonald, T. M. (2010). The Peptide That Binds. *Harvard Review of Psychiatry*, *18*(1), 1–21.

Mazar, N., & Aggarwal, P. (2011). Greasing the Palm: Can Collectivism Promote Bribery? *Psychological Science*, *22*(7), 843–848.

Mazar, N., Amir, O., & Ariely, D. (2008). The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research*, *45*(6), 633–644.

McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences*, *98*(20), 11832–11835.

McDermott, R., Fowler, J. H., & Smirnov, O. (2008). On the Evolutionary Origin of Prospect Theory Preferences. *The Journal of Politics*, *70*(02), 335–350.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: dynamics of willpower. *Psychological review*, *106*(1), 3.

Modahl, C., Green, L. A., Fein, D., Morris, M., Waterhouse, L., Feinstein, C., & Levin, H. (1998). Plasma oxytocin levels in autistic children. *Biological Psychiatry*, *43*(4), 270–277.

Nagel, R. (1995). Unraveling in Guessing Games: An Experimental Study. *The American economic review*, *85*(5), 1313–1326.

Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants With and Without a Sword: Self-Governance is Possible. *The American Political Science Review*, *86*(2), 404.

Pedersen, C. A., Gibson, C. M., Rau, S. W., Salimi, K., Smedley, K. L., Casey, R. L., . . . Penn, D. L. (2011). Intranasal oxytocin reduces psychotic symptoms and improves Theory of Mind and social perception in schizophrenia. *Schizophrenia research*, *132*(1), 50–53.

Raven, J. C. (1965). *Advanced Progressive Matrices*. London, Testmaterial: Lewis.

Rodrigues, S. M., Saslow, L. R., Garcia, N., John, O. P., & Keltner, D. (2009). Oxytocin receptor genetic variation relates to empathy and stress reactivity in humans. *Proceedings of the National Academy of Sciences*, *106*(50), 21437–21441.

Rubinstein, A. (2007). Instinctive and Cognitive Reasoning: A Study of Response Times. *The Economic Journal*, *117*(523), 1243–1259.

Sapienza, P., Zingales, L., & Maestripieri, D. (2009). Gender differences in financial risk aversion and career choices are affected by testosterone. *Proceedings of the National Academy of Sciences*, *106*(36), 15268–15273.

Schubert, S., & Miller, C. T. (2008, December 20). At Siemens, Bribery Was Just a Line Item. *New York Times*. Retrieved from http://www.nytimes.com/2008/12/21/business/worldbusiness/21siemens.html?pagewanted=all&_r=0

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The Constructive, Destructive, and Reconstructive Power of Social Norms. *Psychological Science*, *18*(5), 429–434.

Schulze, L., Lischke, A., Greif, J., Herpertz, S. C., Heinrichs, M., & Domes, G. (2011). Oxytocin increases recognition of masked emotional faces. *Psychoneuroendocrinology*, *36*(9), 1378–1382.

Schwieren, C., & Weichselbaumer, D. (2010). Does competition enhance performance or cheating? A laboratory experiment. *Journal of Economic Psychology*, *31*(3), 241–253.

Seyfarth, R. M., & Cheney, D. L. (2013). Affiliation, empathy, and the origins of theory of mind. *Proceedings of the National Academy of Sciences*, *110*, 10349–10356.

Shalvi, S., & De Dreu, C. K. W. (2014). Oxytocin promotes group-serving dishonesty. *Proceedings of the National Academy of Sciences*, *111*(15), 5503–5507.

Shalvi, S., Eldar, O., & Bereby-Meyer, Y. (2012). Honesty Requires Time (and Lack of Justifications). *Psychological Science*, *23*(10), 1264–1270.

Shamay-Tsoory, S. G., Abu-Akel, A., Palgi, S., Sulieman, R., Fischer-Shofty, M., Levkovitz, Y., & Decety, J. (2013). Giving peace a chance: oxytocin increases empathy to pain in the context of the Israeli–Palestinian conflict. *Psychoneuroendocrinology*, *38*(12), 3139–3144.

Shamay-Tsoory, S. G., Fischer, M., Dvash, J., Harari, H., Perach-Bloom, N., & Levkovitz, Y. (2009). Intranasal administration of oxytocin increases envy and schadenfreude (gloating). *Biological psychiatry*, *66*(9), 864–870.

Shleifer, A. (2004). Does Competition Destroy Ethical Behavior? *The American economic review*, *94*(2), 414–418.

Singer, T., & Fehr, E. (2005). The neuroeconomics of mind reading and empathy. *The American economic review*, *95*(2), 340–345.

Stallen, M., De Dreu, C. K. W., Shalvi, S., Smidts, A., & Sanfey, A. G. (2012). The Herding Hormone: Oxytocin Stimulates In-Group Conformity. *Psychological Science*, *23*(11), 1288–1292.

Sutter, M., Haigner, S., & Kocher, M. G. (2010). Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations. *The Review of Economic Studies*, *77*(4), 1540–1566.

Utikal, V., & Fischbacher, U. (2013). Disadvantageous lies in individual decisions. *Financial Sector Performance and Risk*, *85*(0), 108–111.

van Beest, I., & Williams, K. D. (2006). When inclusion costs and ostracism pays, ostracism still hurts. *Journal of Personality and Social Psychology*, *91*(5), 918–928.

van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2012). A sniff of trust: meta-analysis of the effects of intranasal oxytocin administration on face recognition, trust to in-group, and trust to out-group. *Psychoneuroendocrinology*, *37*(3), 438–443.

Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology*, *51*(1), 110–116.

Zak, P. J., & Knack, S. (2001). Trust and Growth. *The Economic Journal*, *111*(470), 295–321.

Zak, P. J., Kurzban, R., & Matzner, W. T. (2004). The Neurobiology of Trust. *Annals of the New York Academy of Sciences*, *1032*(1), 224–227.

Zak, P. J., Kurzban, R., & Matzner, W. T. (2005). Oxytocin is associated with human trustworthiness. *Hormones and Behavior*, *48*(5), 522–527.

Zak, P. J., Stanton, A. A., Ahmadi, S., & Brosnan, S. (2007). Oxytocin Increases Generosity in Humans. *PLoS ONE*, *2*(11), e1128.