



Trust and Testosterone: An Experimental Design

**Werner Raub
(joint work with
Vincent Buskens & Jack van Honk)**

**Treffen der Sektion Ökonomik und empirische Soziologie
der Leopoldina, Bonn, 18.-19. Februar 2010**



Neuroeconomics: survey of the literature

Journal of Economic Literature
Vol. XLIII (March 2005), pp. 9–64

Neuroeconomics: How Neuroscience Can Inform Economics

COLIN CAMERER, GEORGE LOEWENSTEIN, and DRAZEN PRELEC*

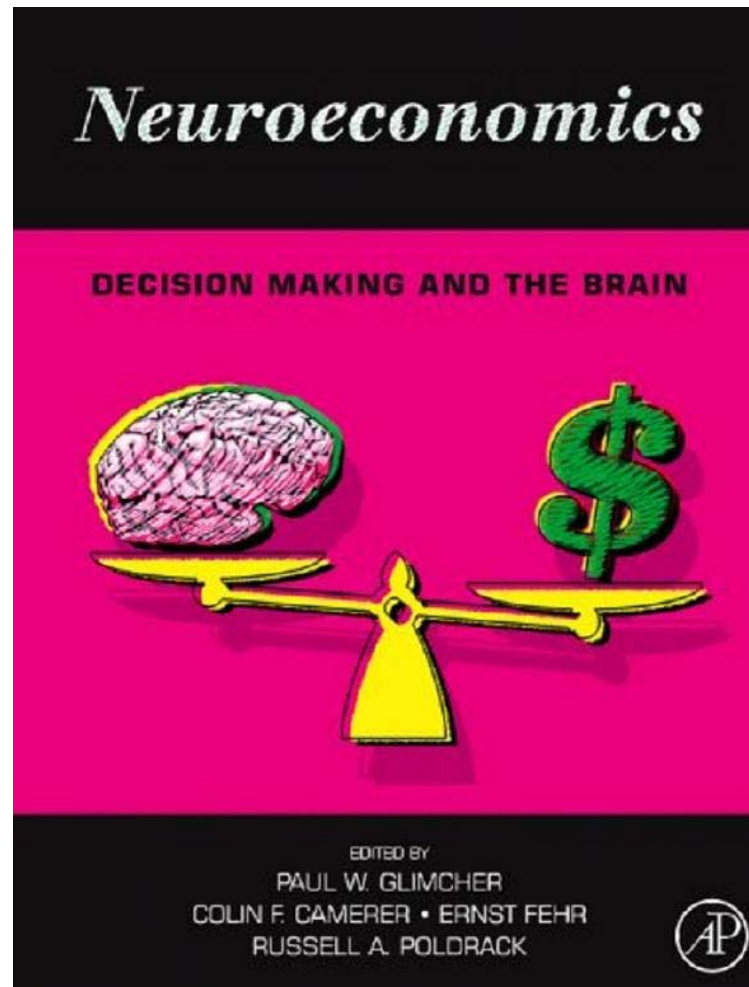
Who knows what I want to do? Who knows what anyone wants to do? How can you be sure about something like that? Isn't it all a question of brain chemistry, signals going back and forth, electrical energy in the cortex? How do you know whether something is really what you want to do or just some kind of nerve impulse in the brain. Some minor little activity takes place somewhere in this unimportant place in one of the brain hemispheres and suddenly I want to go to Montana or I don't want to go to Montana. (*White Noise*, Don DeLillo)

1. Introduction

In the last two decades, following almost a century of separation, economics has begun to import insights from psychology. “Behavioral economics” is now a prominent fixture on the intellectual landscape and has spawned applications to topics in economics,

such as finance, game theory, labor economics, public finance, law, and macroeconomics (see Colin Camerer and George Loewenstein 2004). Behavioral economics has mostly been informed by a branch of psychology called “behavioral decision research,” but other cognitive sciences are ripe for harvest. Some important insights will surely come from neu-

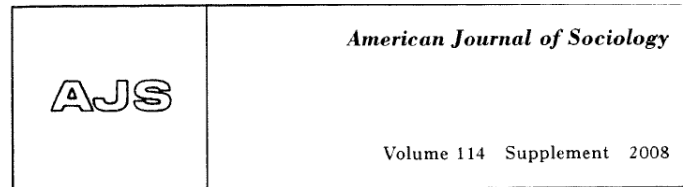
Neuroeconomics: handbook



Criticism...

- ***From a theoretical and philosophy of science perspective: Gul, Faruk & Wolfgang Pesendorfer (2005) The Case for Mindless Economics, mimeo, Princeton University***
- ***On empirical work in neuroeconomics and social neuroscience: Vul, Edward, Christine Harris, Piotr Winkielman & Harold Pashler (2008) Puzzling High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition, Perspectives on Psychological Science 4(3): 274-290 (formerly known as "Voodoo Correlations in Social Neuroscience"), see also comments and reply***

(Somewhat) related research in sociology



EXPLORING GENETICS AND SOCIAL STRUCTURE

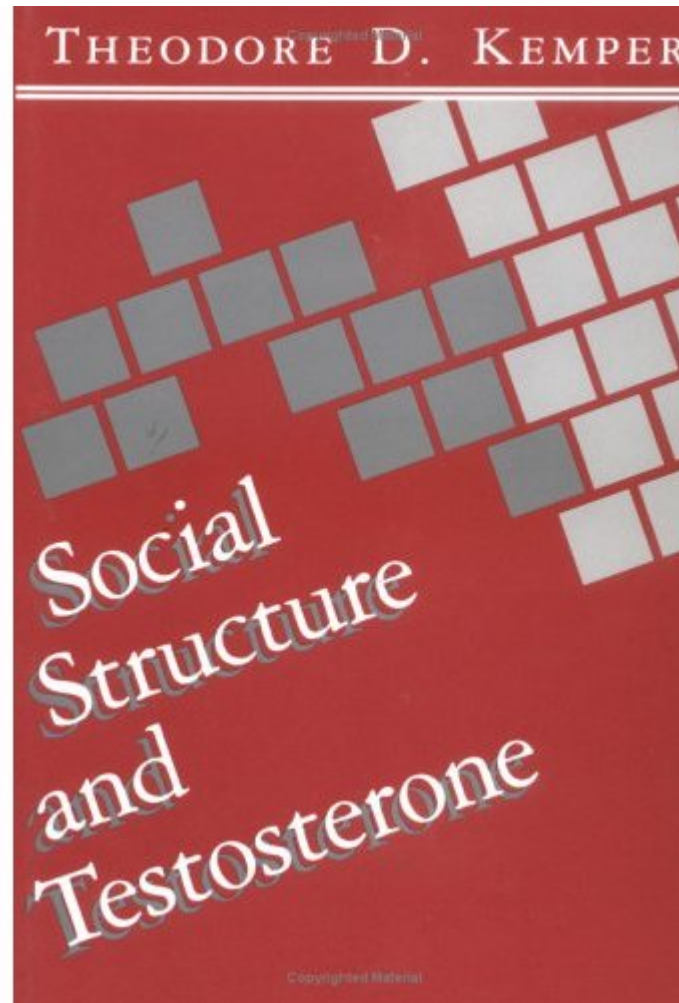
Guest Editor, Peter Bearman

Guest Associate Editors, Molly A. Martin and Sara Shostak

CONTENTS

- v Introduction: Exploring Genetics and Social Structure
PETER BEARMAN
- S1 Genetics and the Social Science Explanation of Individual Outcomes
JEREMY FREESE
- S36 Gene by Social Context Interactions for Number of Sexual Partners among White Male Youths: Genetics-Informed Sociology
GUANG GUO, YUYING TONG, AND TIANJI CAI
- S67 The Intergenerational Correlation in Weight: How Genetic Resemblance Reveals the Social Role of Families
MOLLY A. MARTIN
- S106 Reconstructing Race in Science and Society: Biology Textbooks, 1952–2002
ANN MORNING
- S138 Gender Differences in Extreme Mathematical Achievement: An International Perspective on Biological and Social Factors
ANDREW M. PENNER
- S171 Under the Influence of Genetics: How Transdisciplinarity Leads Us to Rethink Social Pathways to Illness
BERNICE A. PESCOLIDLO, BREA L. PERRY, J. SCOTT LONG, JACK K. MARTIN, JOHN I. NURNBERGER, JR., AND VICTOR HESSELBROCK

Related research in sociology



Neuroscience empirical methods

- **Lesion studies and brain stimulation**
- **fMRI**
- **Drug effects**

etc...

**Overview: Houser & McCabe 2009 in
Glimcher et al. (eds.) (2009)**



Core features of the presentation

- **Aim: applying neuroscience theoretical ideas and empirical methods in the study of social dilemmas – trust problems as an example**
- **Method: drug effects**
 - **Influential previous studies: Kosfeld et al. 2005: effects of oxytocin; Eisenegger et al. 2009: effects of testosterone**
 - **We look at testosterone effects**
- **Limitation: theory, hypotheses, and experimental design only**



Drug effects in social interactions

Vol 435|2 June 2005|doi:10.1038/nature03701

nature

LETTERS

Oxytocin increases trust in humans

Michael Kosfeld^{1*}, Markus Heinrichs^{2*}, Paul J. Zak³, Urs Fischbacher¹ & Ernst Fehr^{1,4}

Trust pervades human societies^{1,2}. Trust is indispensable in friendship, love, families and organizations, and plays a key role in economic exchange and politics³. In the absence of trust among trading partners, market transactions break down. In the absence of trust in a country's institutions and leaders, political legitimacy breaks down. Much recent evidence indicates that trust contributes to economic, political and social success^{4,5}. Little is known, however, about the biological basis of trust among humans. Here we show that intranasal administration of oxytocin, a neuropeptide that plays a key role in social attachment and affiliation in non-human mammals⁶⁻⁸, causes a substantial increase in trust among humans, thereby greatly increasing the benefits from social interactions. We also show that the effect of oxytocin on trust is

monetary payoff. However, the trustee also has the option of violating the investor's trust. As sharing the proceeds is costly for the trustee, a selfish trustee will never honour the investor's trust because the investor and the trustee interact only once during the experiment.

The investor is therefore caught in a dilemma: if he trusts and the trustee shares, the investor increases his payoff, but he is also subject to the risk that the trustee will abuse this trust. In the latter case, the investor is worse off than if he had not trusted at all and, adding insult to injury, the trustee has an unfair payoff advantage relative to the investor. Substantial evidence exists to show that humans are averse to such risks²²⁻²⁴. Moreover, the aversion of investors to abuse of trust seems to have an important role across different human cultures and

Drug effects in social interactions

doi:10.1038/nature08711

nature

LETTERS

Prejudice and truth about the effect of testosterone on human bargaining behaviour

C. Eisenegger¹, M. Naef^{1,2}, R. Snozzi¹, M. Heinrichs³ & E. Fehr¹

Both biosociological and psychological models, as well as animal research, suggest that testosterone has a key role in social interactions^{1–7}. Evidence from animal studies in rodents shows that testosterone causes aggressive behaviour towards conspecifics⁷. Folk wisdom generalizes and adapts these findings to humans, suggesting that testosterone induces antisocial, egoistic, or even aggressive human behaviours. However, many researchers have questioned this folk hypothesis^{1–6}, arguing that testosterone is primarily involved in status-related behaviours in challenging social interactions, but causal evidence that discriminates between these views is sparse. Here we show that the sublingual administration of a single dose of testosterone in women causes a substantial increase in fair bargaining behaviour, thereby reducing bargaining conflicts and increasing the efficiency of social interactions. However, subjects who believed that they received testosterone—regardless of whether they actually received it or not—behaved much more unfairly than those who believed that they were treated with placebo. Thus, the folk hypothesis seems to generate a strong negative association between subjects' beliefs and the fairness of their offers, even though testosterone administration actually

hypothesis, testosterone induces status seeking, in particular in those social situations that constitute a potential challenge to a person's status. Thus, in settings such as prisons, where rigid social hierarchies impose subordinate positions on individuals, those who are predisposed to seek social status may question the hierarchy in antisocial and rebellious ways. The evidence mentioned above^{10,11} is thus also consistent with the social status hypothesis.

Although the social status hypothesis constitutes a plausible alternative to the folk hypothesis, it unfortunately remains largely based on correlative evidence^{1–5}. However, a clean separation of the two hypotheses is possible because testosterone-induced status seeking may take a prosocial dimension if the prosocial behaviour enables individuals to master a challenge in order to secure their social position and thus to attain access to resources. Among the interactive games developed to examine prosocial behaviour^{12,13}, the ultimatum bargaining game^{14–20} can be used for this purpose. In this game real money is at stake and two parties, A and B, have to agree on the division of 10 money units (MUs). Party A, the proposer, can propose how the 10 MUs will be allocated between A and B. Party B, the responder, can only accept or reject A's proposal, but cannot make a



Outline

- How to explain trust (trustfulness and trustworthiness)?
- Competing general hypotheses on testosterone effects on behavior in trust problems
- Design for an experimental test using the Trust Game
- Discussion
 - (Why) is all this relevant for sociology?
 - Some related research
 - Some open questions and issues

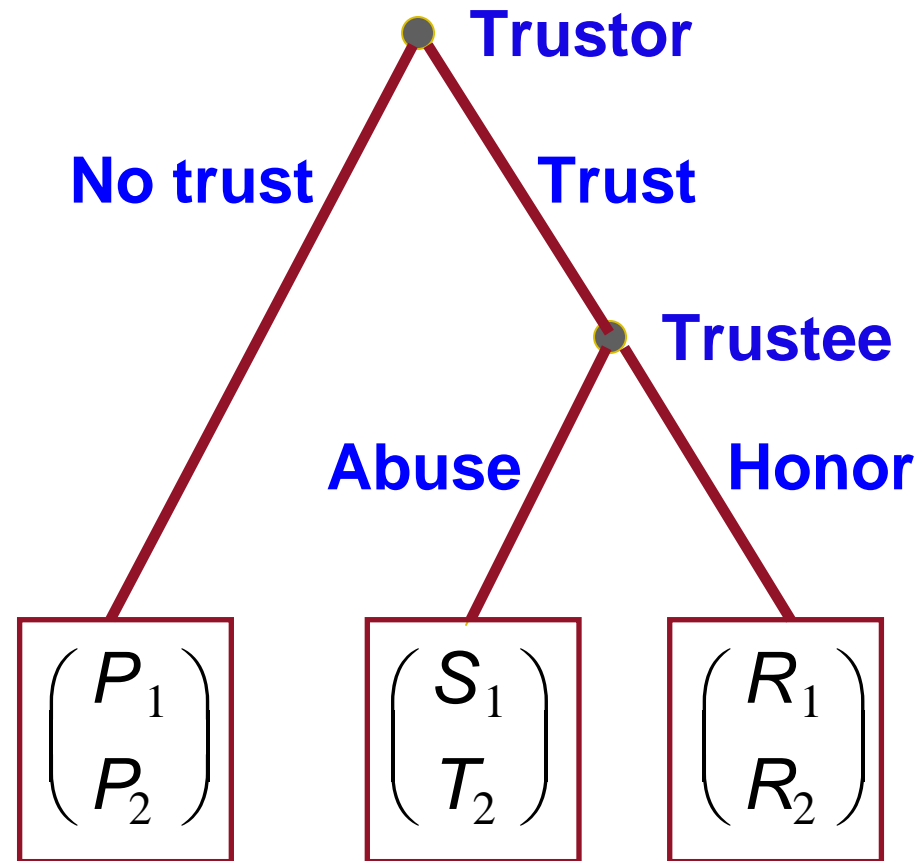


How to explain trust?



Universiteit Utrecht

Trust Game



$$S_1 < P_1 < R_1$$
$$P_2 < R_2 < T_2$$

Standard assumptions and implications

- **A1: Game-theoretic rationality (equilibrium behavior)**
- **A2: Selfishness: “utility = own money”**
- **Implication from A1 and A2:**
 - ***One-shot Trust Game:* no trustfulness, no trustworthiness**
 - ***Repeated Trust Game:* under suitable conditions, trustfulness and trustworthiness (at least for many periods of the repeated game)**



Trust as a result of repeated interactions

- If the Trust Game is repeated, trustfulness and trustworthiness can be a result of conditional cooperation of the trustor and – in the case of a finitely repeated game with incomplete information – reputation building of the trustee
- *Note:* trust as a result of individually rational behavior of selfish actors (“trust as a result of enlightened self-interest”)

Empirical evidence

- ***One-shot Trust Game:*** trustors are sometimes trustful, trustees are sometimes trustworthy
- ***Repeated Trust Game:*** less trustfulness and less trustworthiness than implied by the standard assumptions
- **General impression:** “standard economic models based on full rationality and narrow self-interest tend to overestimate the difficulties of promoting trust in one-shot situations, and underestimate the difficulties in ongoing interactions” (Bolton & Ockenfels 2009: 31)



Alternatives to the standard assumptions

- Trust through bounded rationality (dropping the assumption of game-theoretic rationality)
- Trust through possibly non-selfish trustees (i.e., heterogeneity of trustees with respect to their preferences) and incomplete information of the trustor on the trustee's preferences (dropping the selfishness assumption)



Replacing the standard assumptions: methodological problems

- Replacing the standard assumptions of game-theoretic rationality and selfishness typically leads to less parsimonious and more complex assumptions that are endangered by less testability in the Popperian sense
- Hence, *new predictions* from alternative assumptions are needed
 - This is sometimes neglected in neuroeconomics studies
 - But see Fehr & Camerer 2007 (*Trends in Cognitive Science*) for a similar argument on “out of treatment forecasting”



Aim of our contribution

- **Use neuroscience tools to generate and empirically test new predictions that follow from non-standard assumptions on decision-making in interdependent situations**

Two general hypotheses on testosterone effects

- ***H1***: Increased testosterone levels are associated with behavior that represents increased selfishness and a desire for “eminence” à la Hobbes (“emotional effect”; cf. the “folk hypothesis” of Eisenegger et al. 2009)
- ***H2***: Increased testosterone levels are associated with behavior of actors as if they are more inclined to individual rationality (“cognitive effect”)

Hypotheses on testosterone effects: remarks

- The two hypotheses are *distinct* and *competing* hypotheses: individual rationality and selfishness are not the same.
- The hypotheses are on associations between testosterone levels and observable behavior rather than on underlying psychobiological mechanisms that generate such associations
 - On such underlying mechanisms:
 - Van Honk & Schutter 2007;
 - Eisenegger 2009

Testosterone effects on behavior: some empirical evidence

- **Animal studies (experiments) and humans (correlative studies): testosterone -> aggressive behavior (TG: abuse trust, being easily provokable as trustor)**
- **Humans (correlative studies): testosterone -> striving for dominance, status, leadership (TG: trustworthy behavior not excluded)**
- **Correlative study: testosterone -> higher financial returns for traders in a stock market (Coates & Herbert 2008), possibly indicating a more rational way of managing risks (TG: more farsightedness in repeated TG)**





Approach for an experimental test of the two general hypotheses using the Trust Game



General idea

- Test implications of the two hypotheses for *one-shot* versus *repeated* Trust Games



H1: testosterone

→ selfishness and eminence

- ***Prediction:*** increased testosterone levels are related to less trustfulness as well as less trustworthiness in *both* the *one-shot* and the *repeated* Trust Game
- ***Intuition*** for prediction:
 - ***One-shot game:*** Consider “non-standard” utility models such as the F&S utility function (“inequity aversion”) with guilt and envy parameters. *H1* implies less guilt for trustee and more envy for trustor.
 - ***Repeated game:*** *H1* implies that short-term incentives for not placing trust and for abusing trust increase.

H2: testosterone → rationality

- ***Prediction* for *one-shot* Trust Game: no or negative effect of increased testosterone levels on trustfulness as well as trustworthiness**
- ***Intuition* for prediction:**
 - No effect if trust results from (incomplete information on) non-selfish preferences
 - Negative effect if trust results from bounded rationality

H2: testosterone → rationality

- ***Prediction for repeated Trust Game:*** increased testosterone levels are related to more trustfulness as well as more trustworthiness
- ***Intuition for prediction:*** trust as individually rational equilibrium outcome of the repeated game. *H2* implies less deviations from that outcome.



New predictions

- Note that the hypotheses *do* lead to new predictions on the effects of assumptions that are alternatives to the standard assumptions of game-theoretic rationality and selfish preferences
- Note that these predictions are derived using neuroscience tools



Experimental design



Universiteit Utrecht

Lab experiment

- **Subjects play Trust Games in the lab (ELSE lab UU; z-Tree software)**
- **Anonymous interactions with actual other subjects in the lab**
- **Complete game structure provided in the instructions; no deception**
- **Points earned represent actual money for the subjects**



The experimental game

- Rather than the original Trust Game and in order to facilitate comparison, we use the *simplified version of the Investment Game* from Kosfeld et al. 2005: 4 investment levels for trustor
 - Note: hypotheses and predictions are robust relative to this modification



Design details I: before experiment

- Only female subjects
- Several hours before the experiment:
 1. Measurement of baseline testosterone level
 2. Subjects are randomly assigned to two conditions: administered with testosterone versus with placebo (*between subjects*); subjects do not know in what condition they are



Design details II: experiment

- Immediately before experiment: measurement of testosterone levels
- Each subject plays a Dictator Game (measurement of preferences)
- Each subject plays a series of *6 one-shot* Trust Games (stranger matching), each with a different partner, and *1 repeated* Trust Game (6 rounds; partner matching), again with a different partner (*within subjects*)



Design details III: further set-up

- Each subject plays always in the *same role*, either as trustor or as trustee
- Each subject knows that all partners are in the *same* experimental condition (testosterone versus placebo) as the subject herself
- Balanced observations of subjects starting with the one-shot games or starting with the repeated game
- Between one-shot and repeated games: risk task
- Questionnaire at the end: subject characteristics; statements on trust; subjective beliefs on experimental condition (testosterone versus placebo) and beliefs on testosterone effects,...





Discussion



Universiteit Utrecht

(Why) is all this relevant for sociology?

- **Trust Game – social dilemmas – problem of social order**
- **One-shot versus repeated Trust Game: effects of embeddedness on trust – *social* conditions for voluntary cooperation**
- **Additional predictions**
 - **See Raub, De Haan, Buskens & Aleman 2004 for related experimental designs employing fMRI methods**

(Why) is all this relevant for sociology? Continued...

- Hypotheses focus on how testosterone levels ("biology") *interact* with social conditions ("sociology") in their effects on trust: this is a stronger case for including neuroscience arguments and methods in sociological research than exclusively hypotheses on main effects of testosterone levels ("biology") on voluntary cooperation *in addition to* main effects of social conditions ("sociology")
- Analysis can be extended to interaction of testosterone with network embeddedness



Some related research



Universiteit Utrecht

A recent paper in *PNAS* 2009

A randomized trial of the effect of estrogen and testosterone on economic behavior

Niklas Zethraeus^a, Ljiljana Kocoska-Maras^b, Tore Ellingsen^a, Bo von Schoultz^b, Angelica Lindén Hirschberg^b, and Magnus Johannesson^{a,1}

^aDepartment of Economics, Stockholm School of Economics, Box 6501, SE-113 83 Stockholm, Sweden; and ^bDepartment of Woman and Child Health, Division of Obstetrics and Gynecology, Karolinska Institutet, SE-171 77 Stockholm, Sweden

Edited by George A. Akerlof, University of California, Berkeley, CA, and approved February 17, 2009 (received for review December 15, 2008)

Existing correlative evidence suggests that sex hormones may affect economic behavior such as risk taking and reciprocal fairness. To test this hypothesis we conducted a double-blind randomized study. Two-hundred healthy postmenopausal women aged 50–65 years were randomly allocated to 4 weeks of treatment with estrogen, testosterone, or placebo. At the end of the treatment period, the subjects participated in a series of economic experiments that measure altruism, reciprocal fairness, trust, trustworthiness, and risk attitudes. There was no significant effect of estrogen or testosterone on any of the studied behaviors.

sex hormones | trust game | ultimatum game | risk aversion

Humans display sizeable individual variation in economic behaviors. Heterogeneity is large both in the domain of

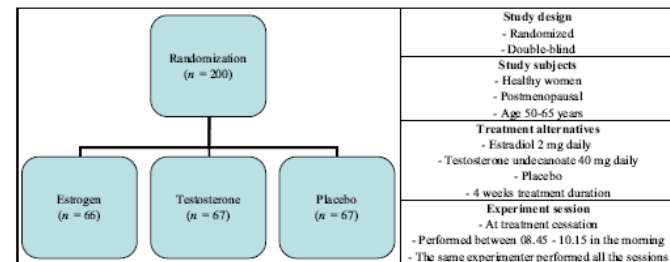


Fig. 1. An economic experiment based on a double-blind randomized trial. Two hundred subjects completed the study; 3 subjects initially randomly assigned did not complete the study (2 in the estrogen group and 1 in the placebo group).

A tiny (and problematic) bit of available empirical evidence

- Zethraeus et al. 2009
- No testosterone effect in one-shot Trust Game
- This is consistent with *H2* (“cognitive effects”) but not with *H1* (“emotional effects”)
- But: design of the study is problematic in various respects
- And: no data for repeated Trust Game



Once again Eisenegger et al. 2009

doi:10.1038/nature08711

nature

LETTERS

Prejudice and truth about the effect of testosterone on human bargaining behaviour

C. Eisenegger¹, M. Naef^{1,2}, R. Snozzi¹, M. Heinrichs³ & E. Fehr¹

Both biosociological and psychological models, as well as animal research, suggest that testosterone has a key role in social interactions^{1–7}. Evidence from animal studies in rodents shows that testosterone causes aggressive behaviour towards conspecifics⁷. Folk wisdom generalizes and adapts these findings to humans, suggesting that testosterone induces antisocial, egoistic, or even aggressive human behaviours. However, many researchers have questioned this folk hypothesis^{1–6}, arguing that testosterone is primarily involved in status-related behaviours in challenging social interactions, but causal evidence that discriminates between these views is sparse. Here we show that the sublingual administration of a single dose of testosterone in women causes a substantial increase in fair bargaining behaviour, thereby reducing bargaining conflicts and increasing the efficiency of social interactions. However, subjects who believed that they received testosterone—regardless of whether they actually received it or not—behaved much more unfairly than those who believed that they were treated with placebo. Thus, the folk hypothesis seems to generate a strong negative association between subjects' beliefs and the fairness of their offers, even though testosterone administration actually

hypothesis, testosterone induces status seeking, in particular in those social situations that constitute a potential challenge to a person's status. Thus, in settings such as prisons, where rigid social hierarchies impose subordinate positions on individuals, those who are predisposed to seek social status may question the hierarchy in antisocial and rebellious ways. The evidence mentioned above^{10,11} is thus also consistent with the social status hypothesis.

Although the social status hypothesis constitutes a plausible alternative to the folk hypothesis, it unfortunately remains largely based on correlative evidence^{1–5}. However, a clean separation of the two hypotheses is possible because testosterone-induced status seeking may take a prosocial dimension if the prosocial behaviour enables individuals to master a challenge in order to secure their social position and thus to attain access to resources. Among the interactive games developed to examine prosocial behaviour^{12,13}, the ultimatum bargaining game^{14–20} can be used for this purpose. In this game real money is at stake and two parties, A and B, have to agree on the division of 10 money units (MUs). Party A, the proposer, can propose how the 10 MUs will be allocated between A and B. Party B, the responder, can only accept or reject A's proposal, but cannot make a



Testosterone effects on bargaining behavior

- Eisenegger et al. 2009 study testosterone effects on behavior in the Ultimatum Game.
- Empirical findings:
 - Proposer: testosterone -> higher offers
 - Responder: no effect of testosterone on rejection rates
 - 'Belief effect': proposers who *believe* that they received testosterone -> *lower* offers
- "Status hypothesis" as explanation: testosterone -> more concern for social status
 - Rejected offers threaten the proposer's status. Low offers can be expected to be often rejected. Hence, testosterone should be related to higher offers.

Implications of the status hypothesis for Trust Game

- Testosterone -> more trustworthiness in repeated TG
- No effect of testosterone on trustworthiness in one-shot TG
- Implications for trustor behavior depend on trustor's expectations of trustee behavior




Implications of our hypotheses for the Ultimatum Game

- **“Emotional effect”**: testosterone -> lower offers; possibly also higher rejection rates
- **“Cognitive effect”**: testosterone -> higher offers (if proposers expect rejection of low offers); effects (if any) on responder behavior depend on responder’s utility function.
 - Note a crucial implicit assumption in the reasoning based on the **“status hypothesis”**: subjects in the proposer role expect that low offers will be rejected. However, under this assumption a rational proposer would likewise make higher/fairer offers. Then, however, we cannot disentangle the **“status effect”** and the **“rationality effect”** of testosterone for the proposer.

Some open issues

- How robust are the implications from the two general hypotheses on testosterone effects also for *other non-standard utility functions* than F&S or ERC?
- Until now, we by-passed a thorough analysis of issues such as “what are actors’ expectations concerning the other actor’s rationality and preferences?”
- We have not yet addressed possible testosterone effects on *expectations concerning the other actor*.
- Testosterone effects on *both selfishness and rationality* would constitute a problem... (intuition for solution: consider TG with *asymmetric* payoffs P_1, P_2 and R_1, R_2)
- We have not yet addressed the possibly “reciprocal” relationship between testosterone and behavior: testosterone -> behavior -> testosterone. This would further strengthen the sociological relevance of research in this field.



- 
- Thanks for your attention!
 - Werner Raub (2009) A Note on Trust and Testosterone, pp. 469-480 in Georg Kamp & Felix Thiele (eds.), *Erkennen und Handeln. Festschrift für Carl Friedrich Gethmann zum 65. Geburtstag*, München: Fink

www.fss.uu.nl/soc/iscore

