ICS
RUG/UU/RU

Trust and Testosterone: A Contribution to Neurosociology

Werner Raub, Vincent Buskens & Jack van Honk

Rational Choice Sociology: Theory and Empirical Applications

Workshop at Venice International University, San Servolo, November 30 – December 4, 2009

venice_09_291109

Universiteit Utrecht

A pioneering study: McCabe et al., *PNAS* 2001

A functional imaging study of cooperation in two-person reciprocal exchange

Kevin McCabe*†‡, Daniel Houser*†§, Lee Ryan*¶, Vernon Smith*†, and Theodore Trouard*‖

*Cognition and Neuroimaging Laboratories, University of Arizona, Tucson, AZ 85721; †Interdisciplinary Center for Economic Science, George Mason University, 4400 University Drive, MSN 1B2, Fairfax, VA 22030; §Department of Economics, McClelland Hall 401, P.O. Box 210108, University of Arizona, Tucson, AZ 85721-0108; ¶Department of Psychology, Psychology 312, P.O. Box 210068, University of Arizona, Tucson, AZ 85721-0068; and ‖Biomedical Engineering Program, AHSC 5302, P.O. Box 245084, Tucson, AZ 85724

Contributed by Vernon Smith, August 7, 2001

Cooperation between individuals requires the ability to infer each other's mental states to form shared expectations over mutual gains and make cooperative choices that realize these gains. From evidence that the ability for mental state attribution involves the use of prefrontal cortex, we hypothesize that this area is involved in integrating theory-of-mind processing with cooperative actions. We report data from a functional MRI experiment designed to test this hypothesis. Subjects in a scanner played standard two-person "trust and reciprocity" games with both human and computer counterparts for cash rewards. Behavioral data shows that seven subjects consistently attempted cooperation with their human counterpart. Within this group prefrontal regions are more active when subjects are playing a human than when they are playing a computer following a fixed (and known) probabilistic strategy. Within the group of five noncooperators, there are no significant differences in prefrontal activation between computer and human conditions.

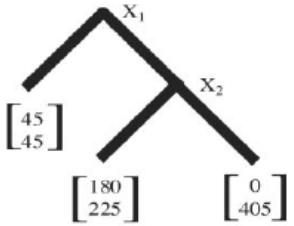


Fig. 1. Diagram of trust game used in the decision making. In the trust game, DM1 moves first (at node x_1) by either moving left, and ending the game, or moving right, giving DM2 a move. If DM1 moves right, DM2 gets the opportunity to move (at node x_2). Once DM2 moves, the game ends, DM1 is paid the top number as a payoff, and DM2 is paid the bottom number as a payoff. By moving right DM1 is trusting DM2 to reciprocate and not defect (move right). By substituting different payoff numbers, different incentives for cooperation can be studied.

Reciprocal exchange (1, 2) is ubiquitous to the behavior of many species (3–5). To make an exchange, it is necessary to overcome the desire for immediate gratification in favor of greater but postponed gains from mutual cooperation. Increased specialization by humans in productive activities, together with the advantages this has produced, likely has been built on improved adaptations for social exchange. The social brain hypothesis (6) explains brain growth as largely an adaptation to

Behavioral Protocol. Subjects responded to cash-payoff salient features of a visually presented two-person binary game tree by pressing response buttons with their right (move right) or left hand (move left). The subjects played the role of either first decision maker or second decision maker in each game. Second decision makers saw the first decision makers' choice before

Universiteit Utrecht

ICS
RUG/UU/RU

Another pioneering study: Sanfey et al., *Science* 2003

and GluR4c are indeed phosphorylated by PKC or if upstream sequence differences and differential protein binding render these subunits incapable of supporting LTD.

Previous attempts to test the involvement of cerebellar LTD in motor learning paradigms have relied on drugs or genetic manipulations that act early in the LTD induction signaling cascade, either at receptors or second messengers (*1*). These studies have been limited owing to the nonspecific nature of the manipulations (e.g., disruption of mGluR1 or PKC function). A GluR2 K882A knock-in mouse could provide the first strong test of the hypothesis that cerebellar LTD is required for certain forms of motor learning.

References and Notes

1. M. F. Bear, D. J. Linden, in *The Synapse*, W. M. Cowan, T. Sudhof, C. F. Stevens, Eds. (Johns Hopkins, Baltimore, MD, 2000), pp. 455–517.
2. C. Hansel, D. J. Linden, E. D'Angelo, *Nature Neurosci.* **4**, 467 (2001).
3. F. Crepel, M. Krupa, *Brain Res.* **458**, 397 (1988).
4. D. J. Linden, J. A. Connor, *Science* **254**, 1656 (1991).
5. N. A. Hartell, *NeuroReport* **5**, 833 (1994).
6. C. I. De Zeeuw *et al.*, *Neuron* **20**, 495 (1998).
7. J. H. Freeman, T. Shi, B. G. Schreurs, *NeuroReport* **9**, 2237 (1998).
8. J. Goossens *et al.*, *J. Neurosci.* **21**, 5813 (2001).
9. M. Ito, M. Sakurai, P. Tongroach, *J. Physiol.* (*London*) **324**, 113 (1982).
10. D. J. Linden, M. H. Dickinson, M. Smeyne, J. A. Connor, *Neuron* **7**, 81 (1991).
11. D. J. Linden, *Learn. Mem.* (*Cold Spring Harbor*) **1**, 121 (1994).
33. A. Nishimune *et al.*, *Neuron* **21**, 87 (1998).

29 January 2003; accepted 6 May 2003

The Neural Basis of Economic Decision-Making in the Ultimatum Game

Alan G. Sanfey,[1,3]* James K. Rilling,[1]* Jessica A. Aronson,[2] Leigh E. Nystrom,[1,2] Jonathan D. Cohen[1,2,4]

The nascent field of neuroeconomics seeks to ground economic decision-making in the biological substrate of the brain. We used functional magnetic resonance imaging of Ultimatum Game players to investigate neural substrates of cognitive and emotional processes involved in economic decision-making. In this game, two players split a sum of money; one player proposes a division and the other can accept or reject this. We scanned players as they responded to fair and unfair proposals. Unfair offers elicited activity in brain areas related to both emotion (anterior insula) and cognition (dorsolateral prefrontal cortex). Further, significantly heightened activity in anterior insula for rejected unfair offers suggests an important role for emotions in decision-making.

Standard economic models of human decision-making (such as utility theory) have typically minimized or ignored the influence of emotions on people's decision-making behavior, idealizing the decision-maker as a perfectly rational cognitive machine. However, in recent years this assumption has been challenged by behavioral economists, who have identified additional

Ultimatum Game. In the Ultimatum Game, two players are given the opportunity to split a sum of money. One player is deemed the proposer and the other, the responder. The proposer makes an offer as to how this money should be split between the two. The second player (the responder) can either accept or reject this offer. If it is accepted, the money is split as proposed,

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-

Universiteit Utrecht

ICS
RUG/UU/RU

Neuroeconomics: special issues of journals

- ***Games and Economic Behavior* 2005**
- ***Economics and Philosophy* 2008**

Papers on decision-making in neuroscience

Glimcher et al. (eds.) 2009, p. 10

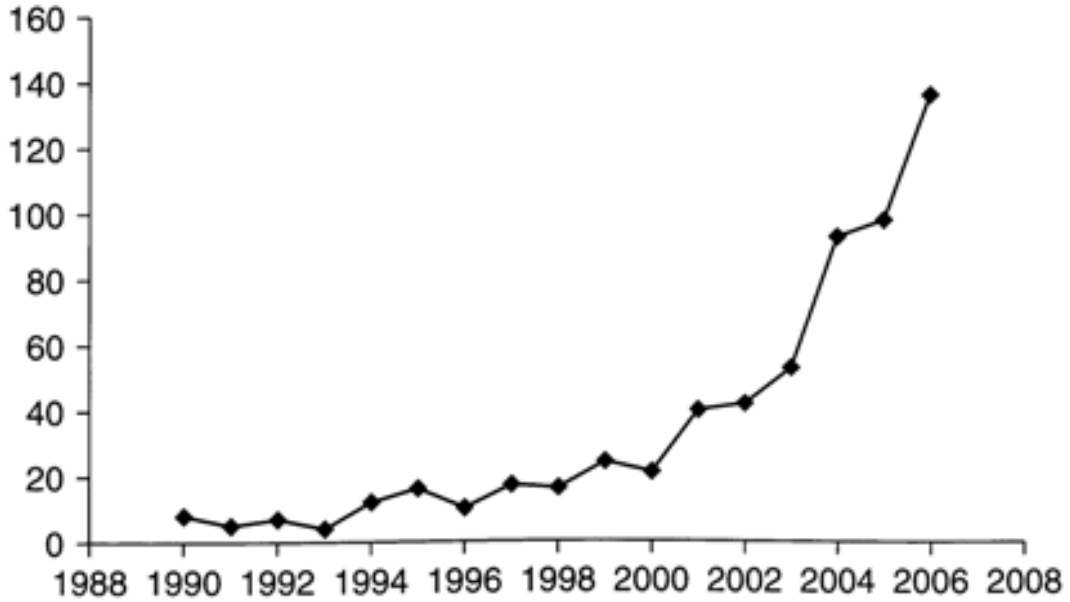
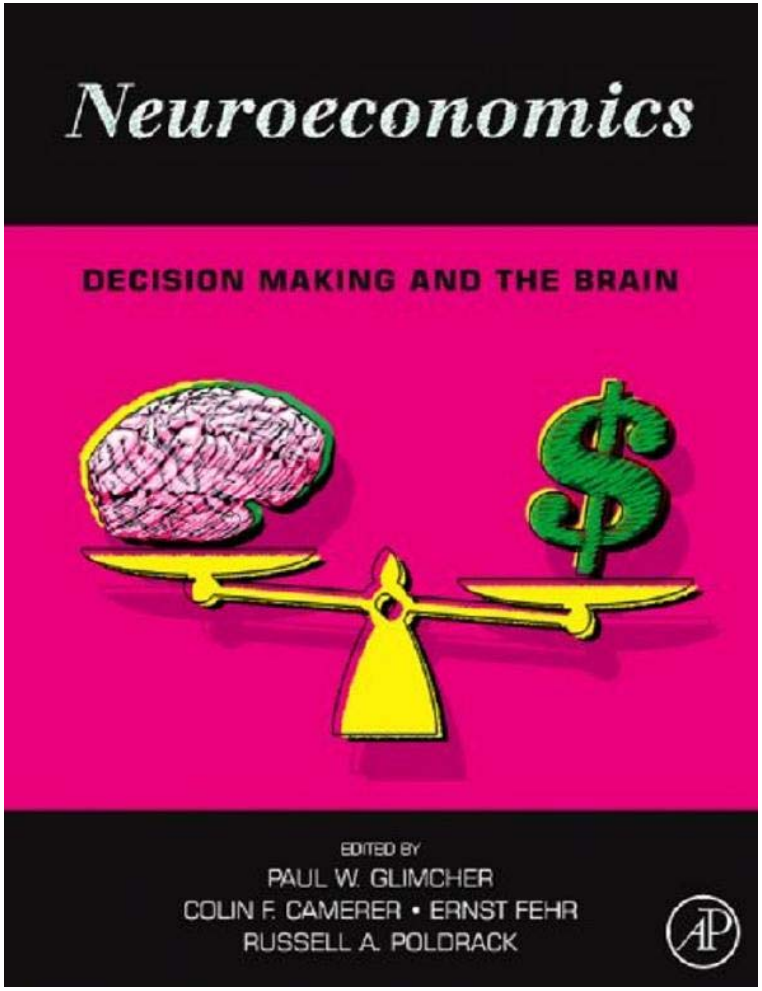


FIGURE 1.1 The increase in numbers of papers on decision-making studies in the neuroscientific literature, 1990–2006

...meanwhile also a handbook



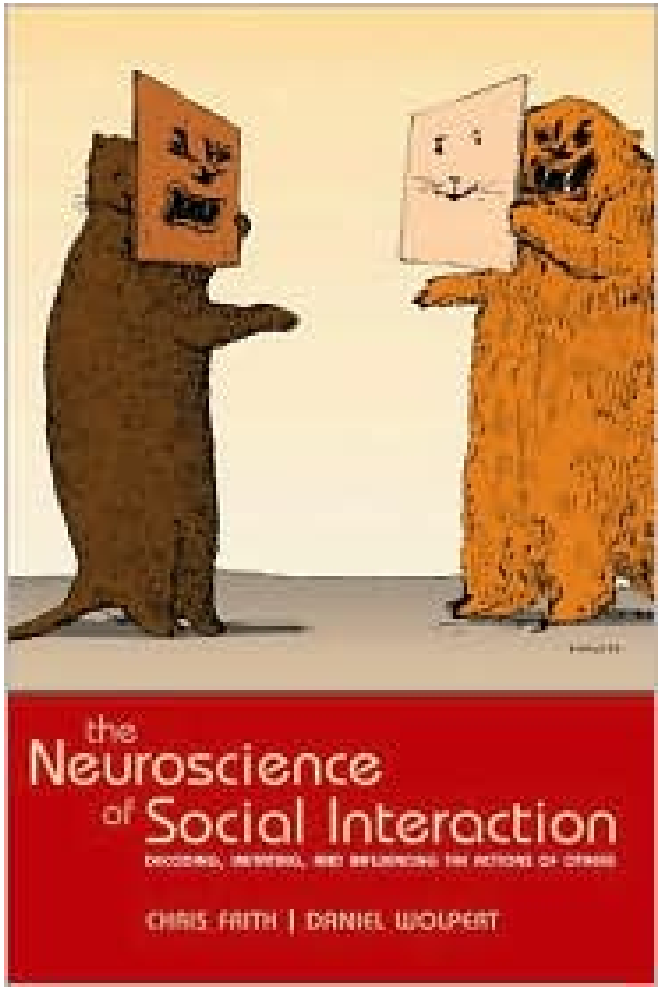
Universiteit Utrecht

ICS
RUG/UU/RU

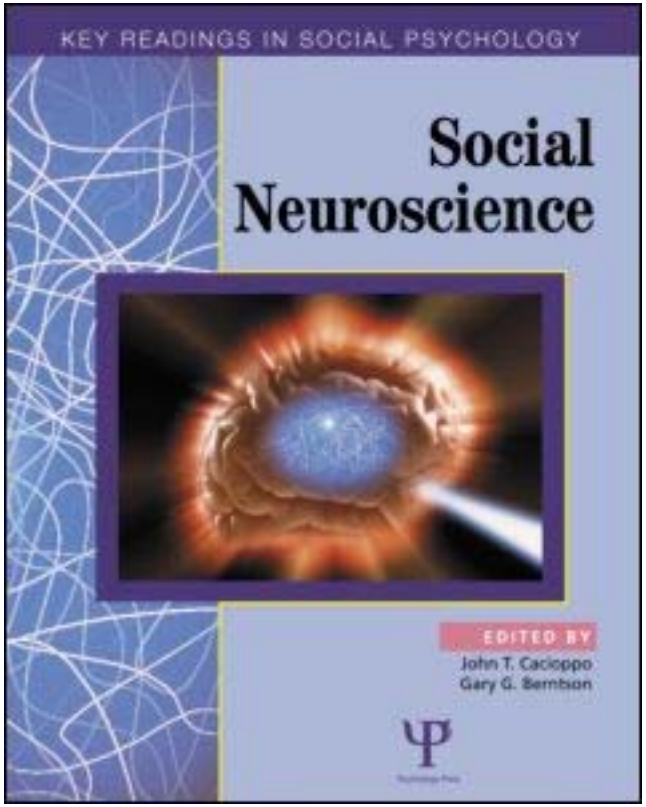
Brief and informative history of the field

- **Glimcher, Paul W., Colin F. Camerer, Ernst Fehr & Russell A. Poldrack (2009) A Brief History of Neuroeconomics, pp. 1-12 in Glimcher et al. (eds.) (2009)**

Social Neuroscience



Social Neuroscience



Universiteit Utrecht

ICS
RUG/UU/RU

And there is 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**

Neuroscience empirical methods

- **Lesion studies**
- **fMRI**
- **Drug effects**

etc…

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

Overview of the presentation

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

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

LETTERS

Oxytocin increases trust in humans

Michael Kosfeld[1]*, Markus Heinrichs[2]*, Paul J. Zak[3], Urs Fischbacher[1] & 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 politics[3]. 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 mammals[6–8], 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.

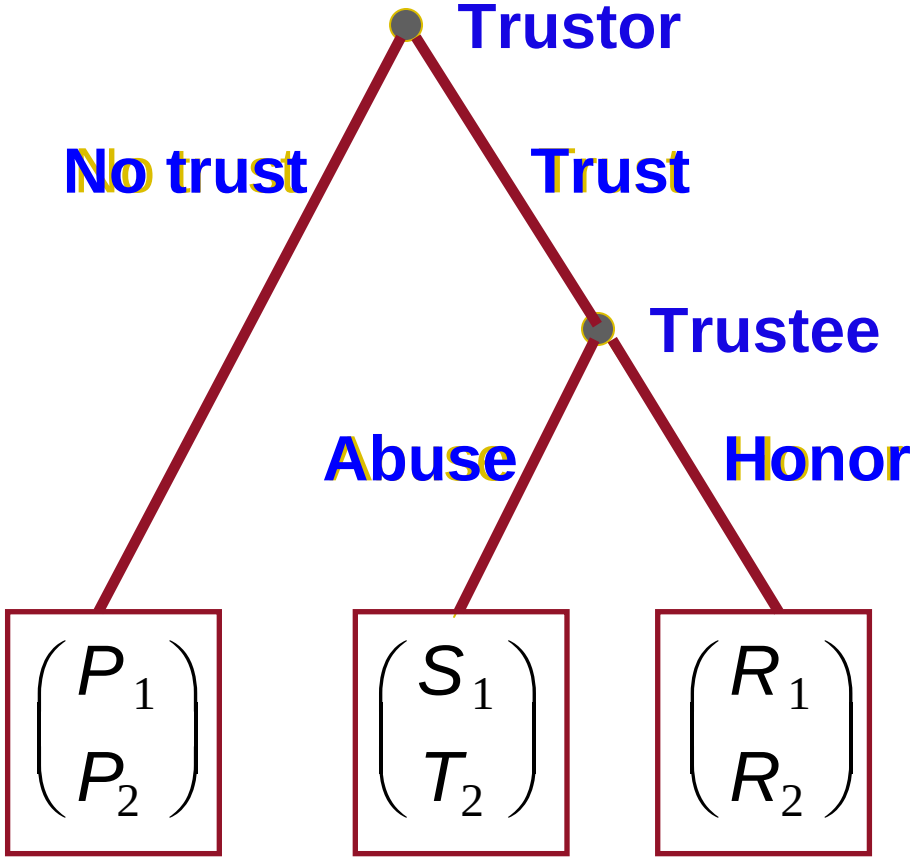
The investor is therefore caught in a dilemma: if he trusts and the trustee shares, the investor increases his payoff, but he is also subject to the risk that the trustee will abuse this trust. In the latter case, the investor is worse off than if he had not trusted at all and, adding insult to injury, the trustee has an unfair payoff advantage relative to the investor. Substantial evidence exists to show that humans are averse to such risks[22–24]. Moreover, the aversion of investors to abuse of trust seems to have an important role across different human cultures and

Outline of the presentation

- How to explain trust (trustfulness and trustworthiness)?
- Competing general hypotheses on testosterone effects on decision making in interdependent situations
- Design for an experimental test using the Trust Game
- Some details of the experimental design
- (Why) is all this relevant for sociology?

How to explain trust?

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 for one-shot Trust Game: no trustfulness, no trustworthiness**
- **However: trustors are sometimes trustful, trustees are sometimes trustworthy**

Trust as a result of repeated interactions

- **If the Trust Game is repeated (indefinitely or, with "some" incomplete information, finitely), 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")**

Trust in one-shot interactions

- **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 often 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 possibly also a desire for "eminence" à la Hobbes ("emotional effect"; cf. "framing" à la Lindenberg)**
- ***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)**

Testosterone effects on behavior: some empirical evidence

- Animal studies: higher testosterone levels often associated with more aggressive behavior (-> TG: abuse trust, being easily provocable as trustor)
- More social species: testosterone related with striving for dominance rather than straightforward aggression. (-> TG: trustworthy behavior not excluded)
  - E.g., studies by Van Honk and colleagues: higher testosterone levels associated with more dominance related personality characteristics
- Higher testosterone levels associated with 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

- ***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 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

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

Universiteit Utrecht

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 conditions 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,...

Conclusion

(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**
- **New predictions**
  - **See Raub, De Haan, Buskens & Aleman 2004 for related experimental design 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 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?**
- **Which *additional* assumptions do we use implicitly in our analysis and, hence, for which *additional* effects (*"alternative explanations"*) do we need to control in the experiment?**
- **Additional measurement of testosterone levels also *after* the experiment?**

A recent paper in *PNAS* 2009

PNAS

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]

[a]Department of Economics, Stockholm School of Economics, Box 6501, SE-113 83 Stockholm, Sweden; and [b]Department 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

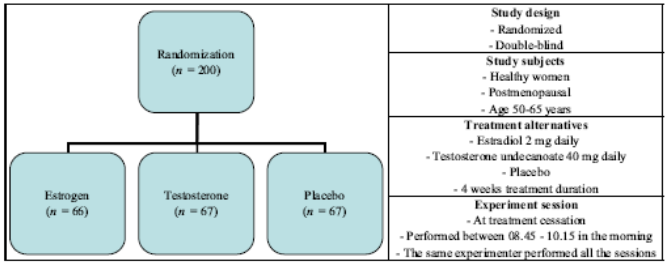


Fig. 1. An economic experiment based on a double-blind randomized trial. Two hundred subjects completed the study; 3 subjects initially randomly assigned did not complete the study (2 in the estrogen group and 1 in the placebo group).

Universiteit Utrecht

ICS
RUG/UU/RU

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**

- **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