Alessandro Innocenti

LabSi
EXPERIMENTAL ECONOMICS LABORATORY
University of Siena

LECTURE 15 TRUST

Aim: To discuss experimental evidence on trust, reciprocity and brain activation.

Outline: Trust game. Trust and reciprocity. Trust and brain activation. Oxytocin and trust. Betrayal aversion.

Readings:

Berg, J., J. Dickhaut, and K. McCabe (1995) "Trust, reciprocity, and social-history", *Games and Economic Behavior*, 10, 122–142.

McCabe, K., D. Houser, L. Ryan, V. Smith, and T. Trouard. (2001) "A functional imaging study of cooperation in two-person reciprocal exchange", *PNAS*, 98, 11832-11835.

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

Blogs, Videos and Websites

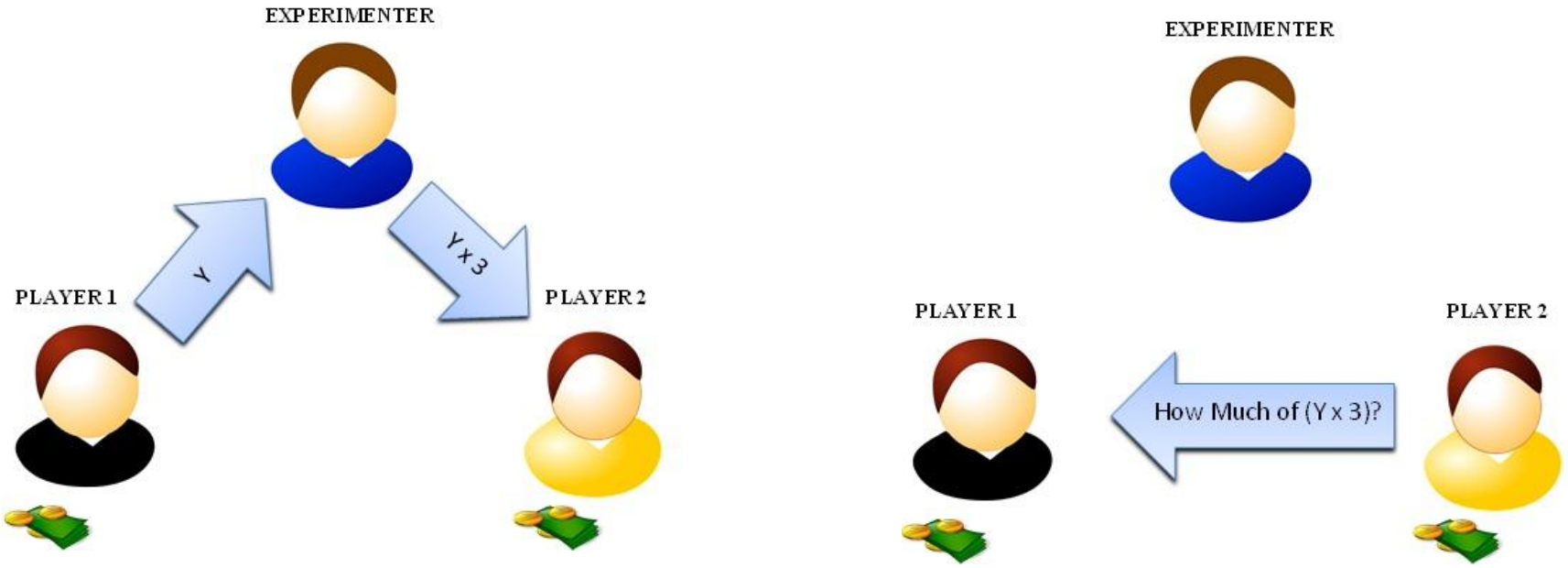
The trust game and wealth creation

http://mindyourdecisions.com/blog/2010/11/02/the-trust-game-and-wealth-creation/#.UhTiKJL0FQC

TRUST GAME

- Two players are paired off anonymously and respectively named as the sender and the responder
- The sender is given a certain amount of money and told that he or she can keep the entire amount or send some or all of it to the responder
- Any money passed from the sender to the responder is tripled by the experimenter and then given to the responder
- The responder can keep the entire amount or give back some or all of it to the sender
- When the sender receives the amount sent back by the responder the game ends

TRUST AND RECIPROCITY



Trust (or investment) game measures

- **propensity to trust** (proportion of the initial endowment sent by the sender)
- **propensity to reciprocate** (ratio between the amount returned and the amount received by the responder)

Backward induction solution:the responder will not send any money back: Anticipating the responder's decision, the sender will not send any money to the responder.

BACKWARD INDUCTION SOLUTION

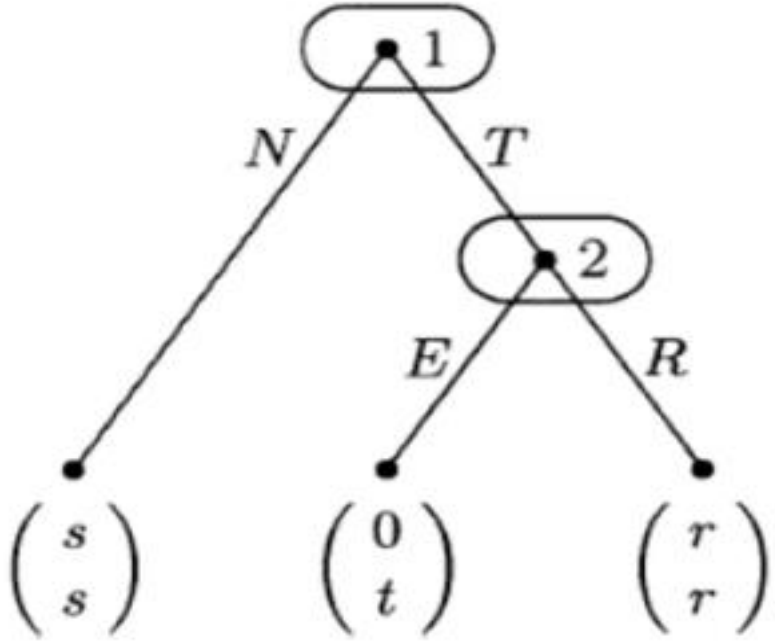


Figure 1: Extensive form of the 'trust' game with t > r > s > 0

Backward induction solution (N.E.)

- the responder will not send any money back
- anticipating the responder's decision, the sender will not send any money to the responder.

FINDINGS

Table 1. Experimental results on the trust game

	Berg-Dickhaut-McCabe (1995)		Buchan-Croson-Johnson (2000)		Schwieren -Sutter (2003)	Burks - Carpenter - Verhoogen (2003)			Chaudhuri - Gangadharan (2003)	Innocenti - Pazienza (2006)
	No history	Social history	Only U.S.	All		Single role	Both roles, no prior	Both roles, prior	Double blind procedure	Double blind procedure
Trust	51.6	53.6	49	67	65.7	65.0	65.2	47.3	43.3	42.2
Reciprocity	30.1	40.2	22	37	37.6	43.6	25.9	17.1	17.5	16.2

Trust = Average fraction sent (Amount sent / Initial endowment); Reciprocity = Average fraction returned (Amount sent back/ Amount received)

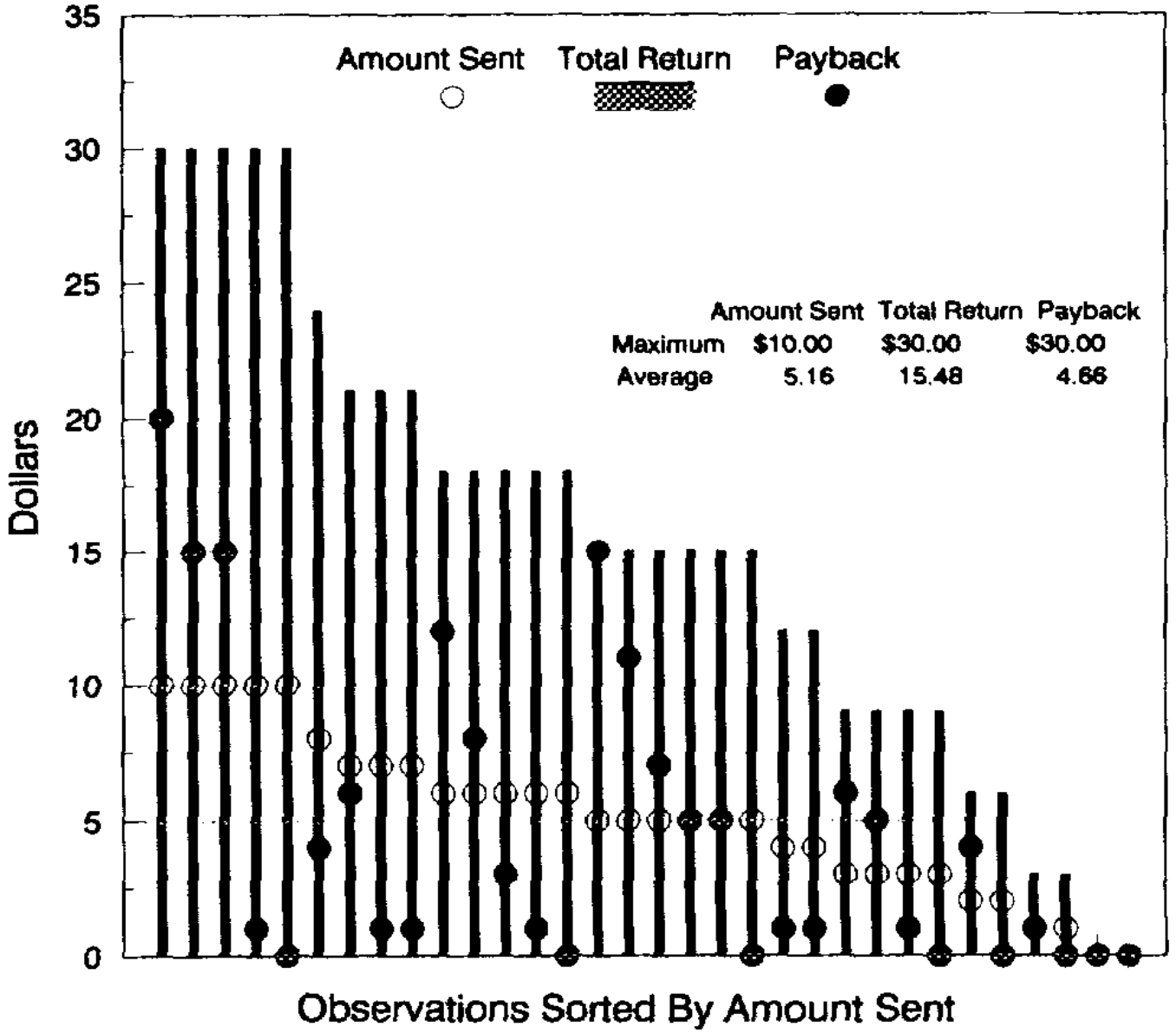
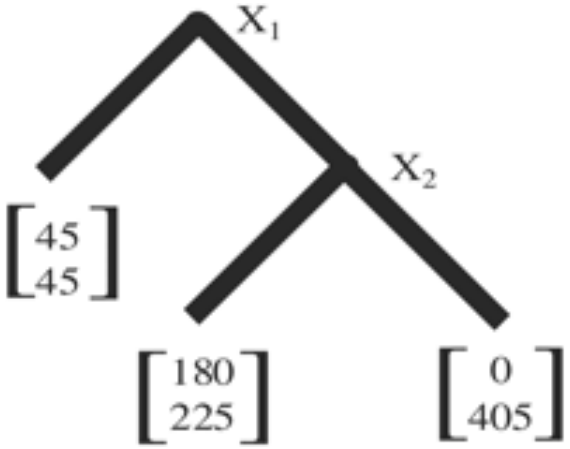


FIG. 2. Trust experiment results showing amount sent (○), total return (■), and payback (●). No history was provided to the subjects.

McCabe et al. 2001

"A functional imaging study of cooperation in two-person reciprocal exchange"
Proc. Nat. Ac. Sci. USA 2001

Design
Subjects play the trust game both against a human opponent and against a computer program which, they were told, would play a human-like strategy.



Cooperation seems to be associated with the activation of the anterior paracingulate cortex, a brain region associated with interpreting and monitoring the mental state of others

DESIGN

Functional magnetic resonance imaging (fMRI)

Data analysis examines the bold response one TR (1.5 s) before the results screen, because decision making for cooperation is likely to be salient at this TR independent of the subject's position in the game

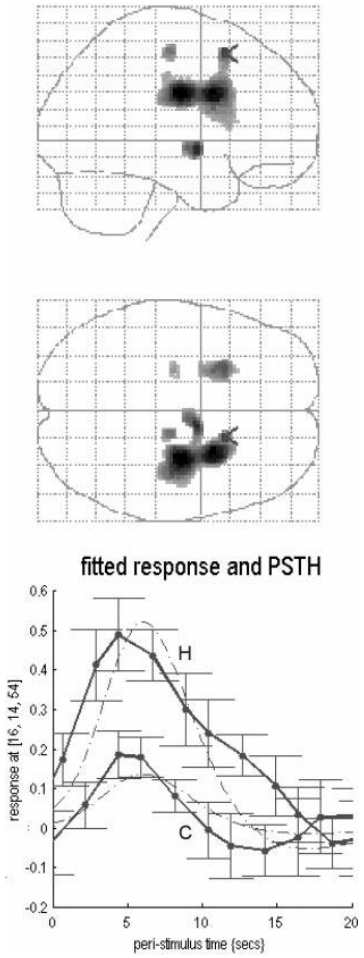
Subjects are likely to ask themselves during this wait condition, "What is my counterpart doing?" and begin to form beliefs about what a delay means about their counterpart's desires.

Hypothesis

Human and computer treatments to generate differential activations associated with predicting and understanding the cooperative intentions of another human.

FINDINGS

- Subjects were more likely to cooperate with real humans than with computers
- Cooperators have a significantly different brain activation in the two conditions (human vs. computer)
- The six subjects with the highest cooperation scores show significant increases in activation in medial prefrontal regions during human-human interactions when compared with human-computer interactions.
- The six subjects who received the lowest cooperation scores (22, 10, 18, 21, 11, and 3) did not show significant activation differences in medial prefrontal cortex between the human and computer conditions.



Bold response of a cooperator for the contrast human (H).computer .
The blobs on the glass brain are clusters of at least 12 contiguous voxels that show significantly more activation in the human than computer condition*. The cursor on the glass brain is* located at the voxel with the greatest *t statistic within the medial prefrontal* clusters. The graph immediately below the glass brains displays the peristimulus time histogram at the voxel indicated by the cursor. This is the mean of the adjusted
(for time and physiological effects) response to the computer and human conditions over all the trials. The bar extends one standard error above and below the mean.

- Behavioural data shows that half the subjects in our experiment **consistently attempted cooperation with their human counterpart.**

- Within this group, and within subjects comparison, they find that **regions of prefrontal cortex** 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 non-cooperators, we find no significant differences in prefrontal cortex between the computer and human conditions.

- One possible explanation for our results is that within this class of games, subjects learn to adopt **game form-dependent rules of thumb** when playing the computer or when playing non-cooperatively with a human counterpart.

- **Cooperation requires an active convergence zone in prefrontal cortex**, that binds joint attention to mutual gains with the inhibition of immediate reward gratification to allow cooperative decisions.

Applications
OXYTOCIN AND TRUSTING BEHAVIOR

Michael Kosfeld, Markus Heinrichs, Paul J. Zak, Urs Fischbacher & Ernst Fehr
"Oxytocin increases trust in humans" *Nature* 2005

In non-human mammals, the neuropeptide oxytocin has a key role in general behavioural regulation, particularly in positive social interactions.

Oxytocin receptors are distributed in various brain regions associated with behaviour, including pair bonding, maternal care, sexual behaviour, and the ability to form normal social attachments.

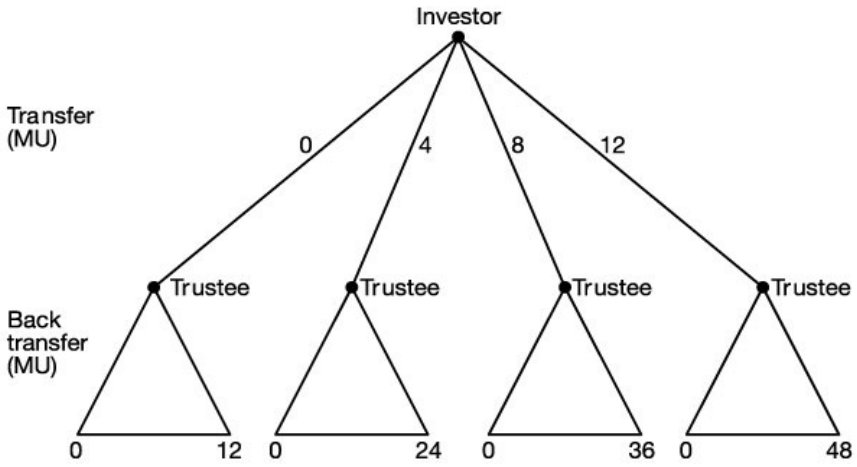
Thus, oxytocin seems to permit animals to facilitate approach behaviour.

HP.: oxytocin might also promote prosocial approach behaviours (such as trust) in humans.

Recent neuroscientific finding: neuropeptides cross the blood-brain barrier after intranasal administration

Double-blind study design to compare:
- trusting behaviour in a group of subjects who receive a dose of intranasal oxytocin
- trusting behaviour in a control group of subjects who receive placebo.

Experiment's object
trust game with real monetary stakes (29+29 subjects)



Hypothesis to test:
oxytocin increases the trusting behaviour of investors $\Rightarrow$ the investors in the oxytocin group will show higher money transfers than those in the placebo group.

Two treatments

- standard trust game
- risk trust game

In B the investor faced the same choices as in the trust game but in which a random mechanism, not the trustee's decision, determined the investor's risk.

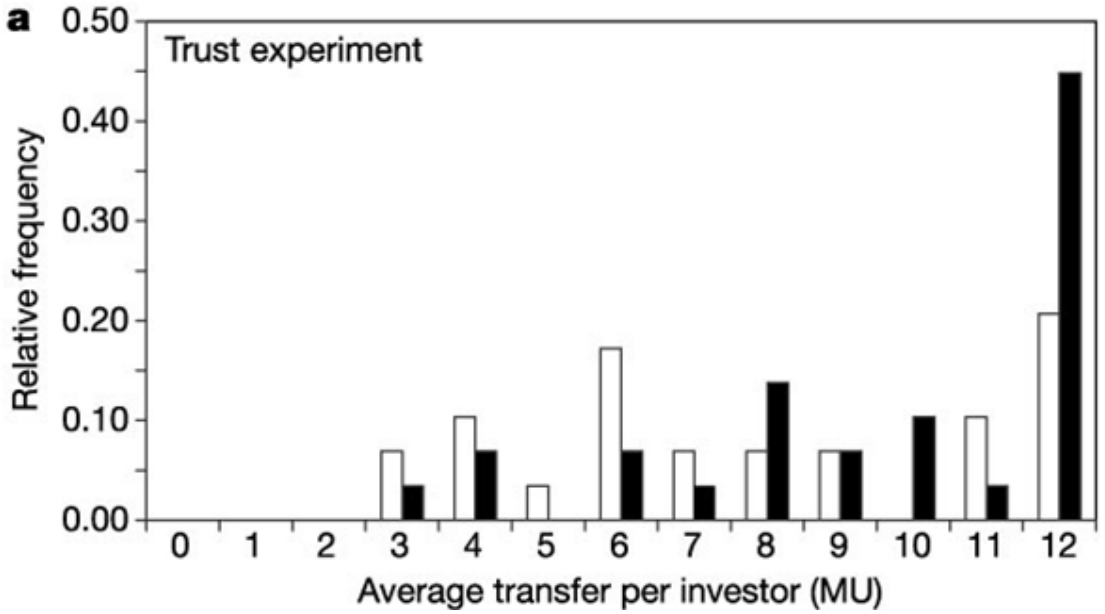
The random mechanism in the risk experiment replicated the trustees' decisions in the trust experiment.

Therefore, the investors faced exactly the same risk as in the trust experiment

However, their transfer decisions were not embedded in a social interaction because there were no trustees in the risk experiment.

Experimental design

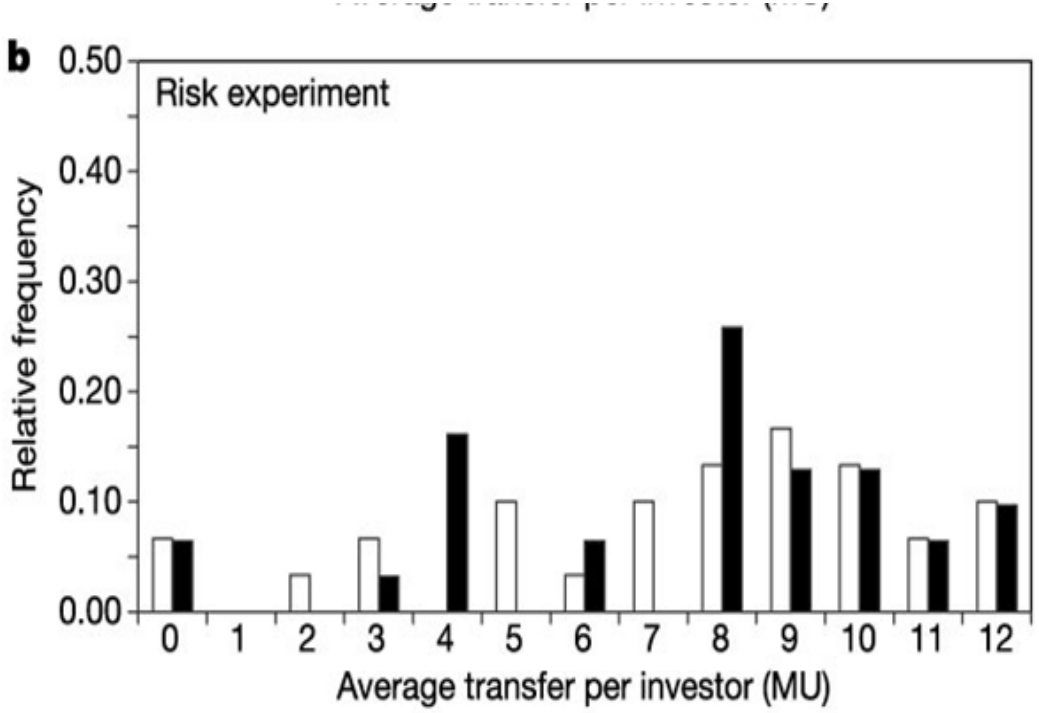
- 194 male students (mean age s.d., 22.0 3.4 yr) from different universities in Zurich
- 128 participants in the trust experiment and 66 subjects participated in the risk experiment
- Exclusion criteria: medical or psychiatric illness, medication, smoking, drug or alcohol abuse
- Subjects were instructed to abstain from food and drink (other than water) for 2 h before the experiment, and from alcohol, smoking and caffeine for 24 h before the experiment
- Participants were informed at the time of recruitment that the experiment would evaluate the effects of a hormone on decision making
- 16 individuals out of the original sample of 194 were excluded because of incorrect substance administration (7 in the trust experiment, 5 in the risk experiment) or their stated disbelief that the opponent in the trust game was actually a human being (4 participants)
- Subjects received a single intranasal dose of 24 IU oxytocin (Syntocinon-Spray, Novartis; 3 puffs per nostril, each with 4 IU oxytocin) or placebo 50 min before the start of the experiment
- Subjects were randomly assigned to the oxytocin or placebo group
- In order to avoid any subjective substance effects (for example, olfactory effects) other than those caused by oxytocin, the placebo contained all inactive ingredients except for the neuropeptide.



Relative frequency of investors' average transfers in oxytocin (filled bars) and placebo (open bars) groups in the trust experiment (n = 58). Subjects given oxytocin show significantly higher transfer levels.

The investors' average transfer is 17% higher in the oxytocin group (Mann-Whitney U-test; z = -1.897, P = 0.029, one-sided).

Median transfer: 10 MU (oxytocin group) > 8 MU (placebo group)



Relative frequency of investors' average transfers in oxytocin (filled bars) and placebo (open bars) groups in the risk experiment (n = 61). Subjects in the oxytocin and the placebo group show statistically identical transfer levels.

Median transfer: 8 MU (in both groups)

Average transfer 7.5 MU (in both groups) (Mann-Whitney U-test; z = 0.022, P = 0.983; two-sided test, n = 31 in oxytocin group, n = 30 in placebo group).

CONCLUSIONS

- oxytocin increases the investors' transfer levels in the trust treatment but not in the risk treatment

- oxytocin specifically affects trust in interpersonal interactions.

INTERPRETATIONS

a) oxytocin causes a general increase in prosocial inclinations

- Oxytocin should affect not only the prosocial behaviour of the investors but also that of the trustees.

- Trustees given oxytocin do not show more trustworthy behaviour.

- At every positive transfer level (4, 8 or 12 MU), their back transfers are statistically indistinguishable from those of placebo trustees (Mann Whitney *U*-tests; $P > 0.243$, two-sided tests for each positive transfer level).

BETRAYAL AVERSION

b) oxytocin does not increase the general inclination to behave prosocially

- Rather, oxytocin specifically affects the trusting behaviour of investors.

c) effect of subjects' beliefs

- Oxytocin might render subjects more optimistic about the likelihood of a good outcome.
- They measured the investor's subjective expectation about the trustee's back transfer after every transfer decision.
- A Mann-Whitney *U*-test indicates that these expectations do not differ significantly between oxytocin and placebo groups at every feasible positive transfer level

d) **oxytocin helps subjects to overcome their betrayal aversion in social interactions**

- This explanation is consistent with the differing effects of oxytocin across the trust and the risk experiments
- It is further supported by the fact that investors faced a considerable betrayal risk.