Lying, ignorance and social dynamics

Prof. Dr. Heiko Rauhut
University of Zurich, Institute of Sociology

March 24, 2020

The preventive effect of ignorance

- Heinrich Popitz (1968): “About the preventive effect of ignorance” [*transl. HR*]

- Counter-intuitive collective phenomenon

  If all norm violations were detected (tax evation, fare-dogding, corruption, moonlighting, adultery, plagiarism etc.), norm violations would spread, norms erode and institutions collapse

- Ignorance hypothesis

  “Veil of ignorance" about norm violations prevents their spread

- Main scope condition

  People underestimate extent of norm violations

The preventive effect of ignorance

Visible power theft triggers its spread (Pakistan/ India)

Maze of illegal connections in Pakistan.
(Business recorder)

Electrical linesman repairs cables of illegal subsidiary wires in India.
(Daily reporter)

The preventive effect of ignorance

Western orientation of TV antennas in GDR and erosion of prohibition norm of Western TV

Abteilung VIII

Dresden, 5. November 1984
Jecz-schl AI/ 3620 /1984

MfS/Dr 43 46
Eing. am 7.11.84
Tgb. Nr. 667
Weiter an:

AG XXII

BStU
000033

Information zu "Turm"

Vor ca. 14 Tagen stellte ein konspirativer Mitarbeiter unserer Diensteinheit fest, daß auf dem Haus

8312 Heidenau,
Beethovenstr. [redacted]

eine Antennenanlage an einem Stahlmast installiert wurde. Diese Anlage besteht aus 2 gekoppelten Antennen, die in Richtung Bayern ausgerichtet sind.

Es wurde festgestellt, daß das Antennenkabel in der Wohnung einer Familie

[redacted]
8312 Heidenau,
Beethovenstr. [redacted]

endet.

Das Haus Beethovenstr. Nr. [redacted] und [redacted] gehören zu einer Wohnzeile (Neubaublock). Auf dem Dach dieser Wohnzeile befindet sich eine Gemeinschaftsantenne für das 1. und 2. DDR-Programm.

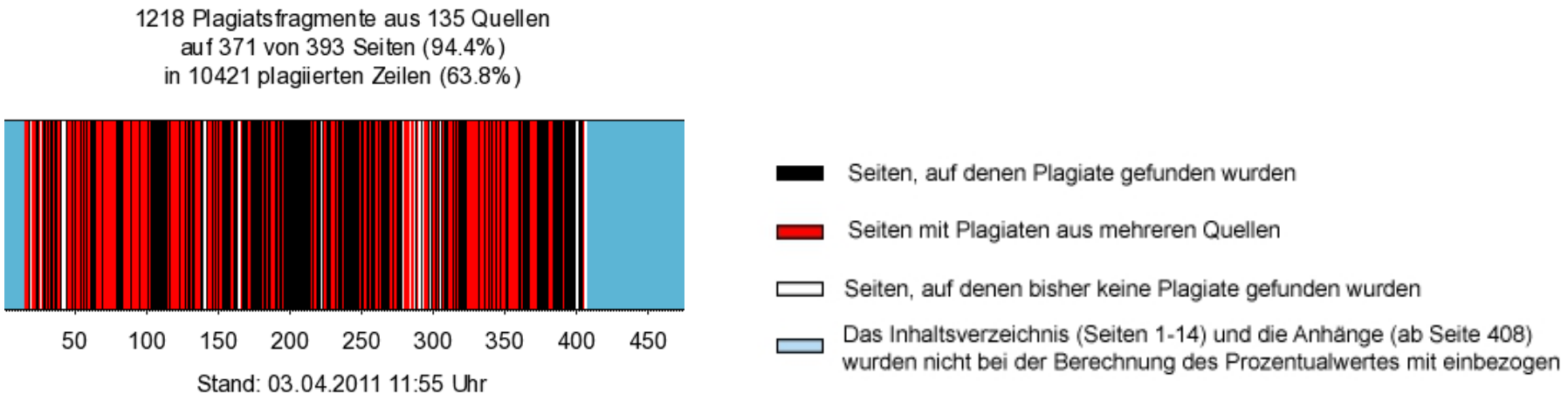
Leiter der Abteilung

Glöckner
Oberst

Plagiarism and its spread

Guttenberg, first statement

Guttenberg, resignation



Broken windows

Wilson & Kelling, 1982, North. Atl.

Pictures: http://herd.typepad.com/, art4chimps.com

Broken windows and cross-norm inhibition

Cross-norm inhibition effect: One norm violation fosters violations of other norms, and disorder spreads from one kind of inappropriate behavior to other kinds.

Violations of the anti-graffiti norm triggers violations of the anti-littering norm

Keizer, Lindenberg, Steg, Science, 2008

Broken windows and cross-norm inhibition

Violations of the anti-bike-parking norm triggers violations of the no-trespassing norm

Fig. 2.

Keizer, Lindenberg, Steg, Science, 2008

Broken windows and cross-norm inhibition

Violations of the shopping-cart-return norm triggers violations of the anti-littering norm (littering a flyer placed at the windshield of parked cars)

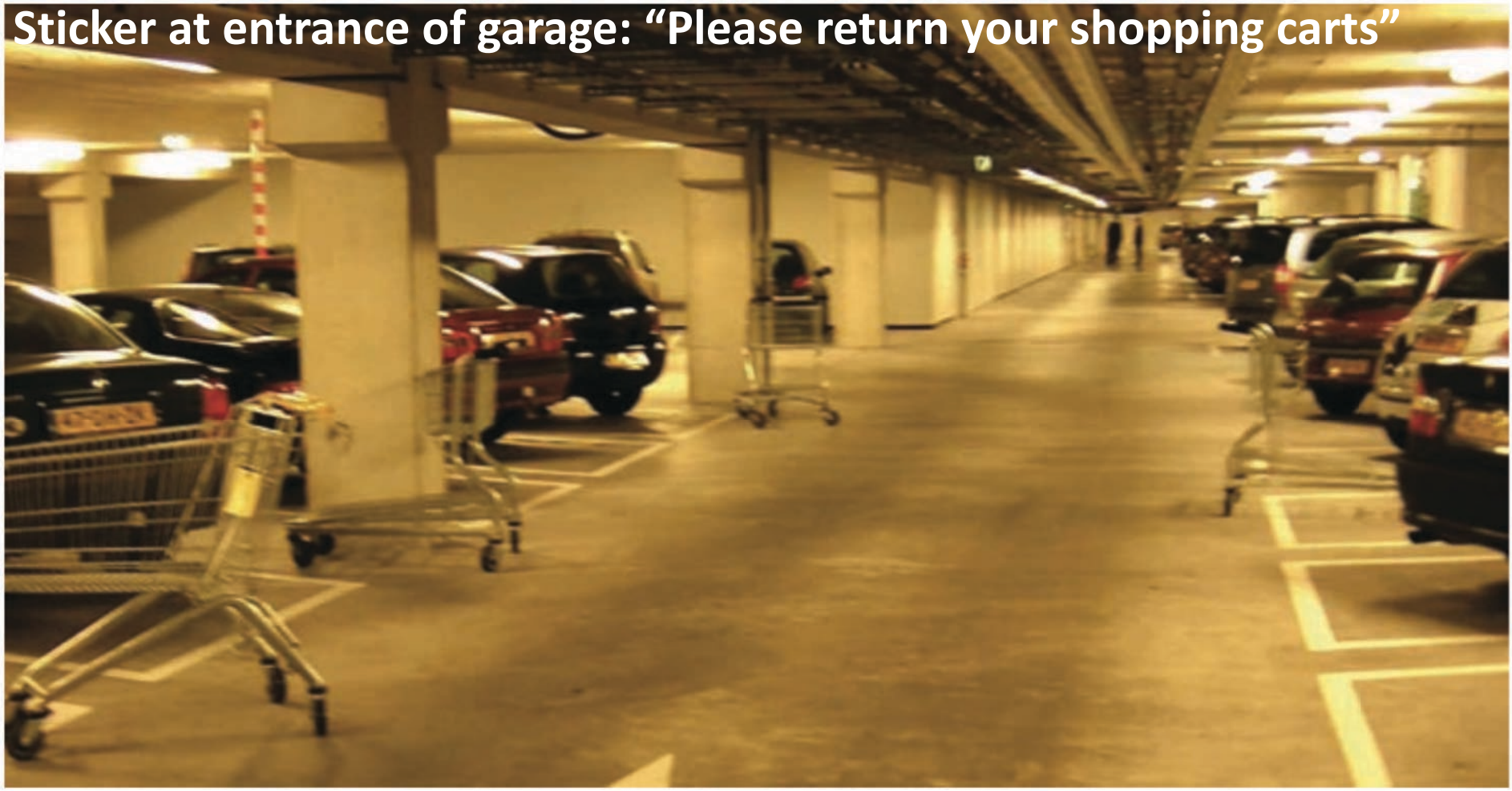


Fig. 3.

Keizer, Lindenberg, Steg, Science, 2008

Broken windows and cross-norm inhibition

Violations of the anti-graffiti norm (the post-box was sprayed with graffiti in the disorder condition) triggers violations of the anti-stealing norm

Fig. 4.

Keizer, Lindenberg, Steg, Science, 2008

The royal lie

Greek philosopher Plato “[. . .] if anyone at all is to have the privilege of lying the rulers of the State [. . .] may be allowed to lie for the public good” (The Royal Lie).

The royal lie

Greek philosopher Plato “[. . .] if anyone at all is to have the privilege of lying the rulers of the State [. . .] may be allowed to lie for the public good” (The Royal Lie).

Hoffmann, Lauer & Rockenbach (2013, JEBO) experimentally show that manipulated feedback of public good provisions can maintain cooperation, in particular, when nobody receives the information of being exploited (i.e. being the sucker)

The royal lie

Greek philosopher Plato “[. . .] if anyone at all is to have the privilege of lying the rulers of the State [. . .] may be allowed to lie for the public good” (The Royal Lie).

Hoffmann, Lauer & Rockenbach (2013, JEBO) experimentally show that manipulated feedback of public good provisions can maintain cooperation, in particular, when nobody receives the information of being exploited (i.e. being the sucker)

Phase 1 (periods 1-20): Information that feedback about others‘ contributions might deviate from actual contributions

Phase 2 (periods 21-40): Similar feedback condition than phase 1, but informed about kind of feedback

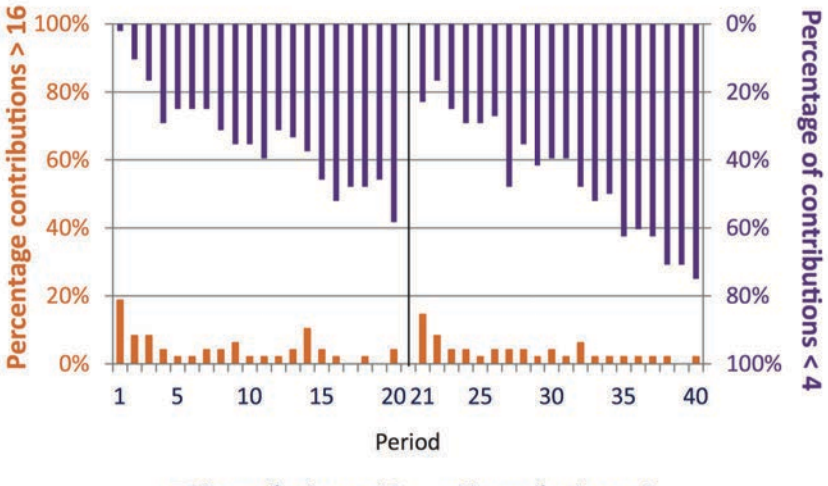
Between phases: Summary feedback about all periods and about whether information feedback deviated or not. Information that phase 2 will be in similar feedback condition than phase 1.

The royal lie

Greek philosopher Plato "[. . .] if anyone at all is to have the privilege of lying the rulers of the State [. . .] may be allowed to lie for the public good" (The Royal Lie).

Hoffmann, Lauer & Rockenbach (2013, JEBO) experimentally show that manipulated feedback of public good provisions can maintain cooperation, in particular, when nobody receives the information of being exploited (i.e. being the sucker)

No royal lie



Phase 1 (periods 1-20): Information that feedback about others' contributions might deviate from actual contributions

Phase 2 (periods 21.40): Similar feedback condition than phase 1, but informed about kind of feedback

Between phases: Summary feedback about all periods and about whether information feedback deviated or not. Information that phase 2 will be in similar feedback condition than phase 1.

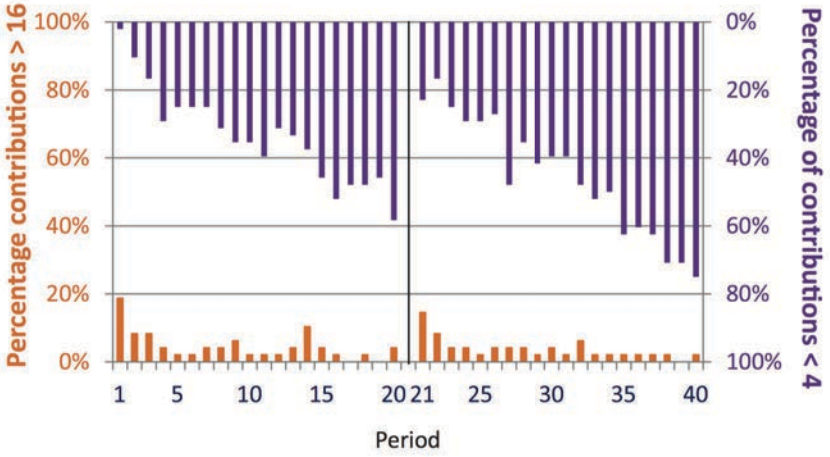
Fig. 4. Development of the percentage of low (<4) and high (>16) contributions.

The royal lie

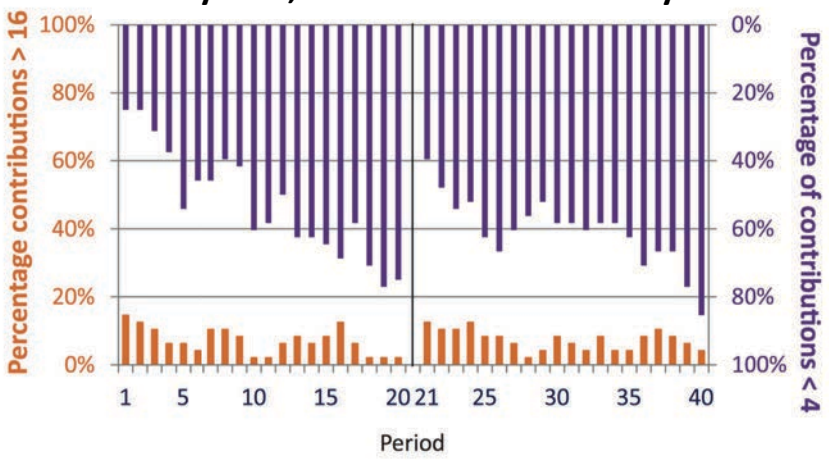
Greek philosopher Plato "[. . .] if anyone at all is to have the privilege of lying the rulers of the State [. . .] may be allowed to lie for the public good" (The Royal Lie).

Hoffmann, Lauer & Rockenbach (2013, JEBO) experimentally show that manipulated feedback of public good provisions can maintain cooperation, in particular, when nobody receives the information of being exploited (i.e. being the sucker)

No royal lie



No royal lie, information uncertainty



Phase 1 (periods 1-20): Information that feedback about others' contributions might deviate from actual contributions

Phase 2 (periods 21.40): Similar feedback condition than phase 1, but informed about kind of feedback

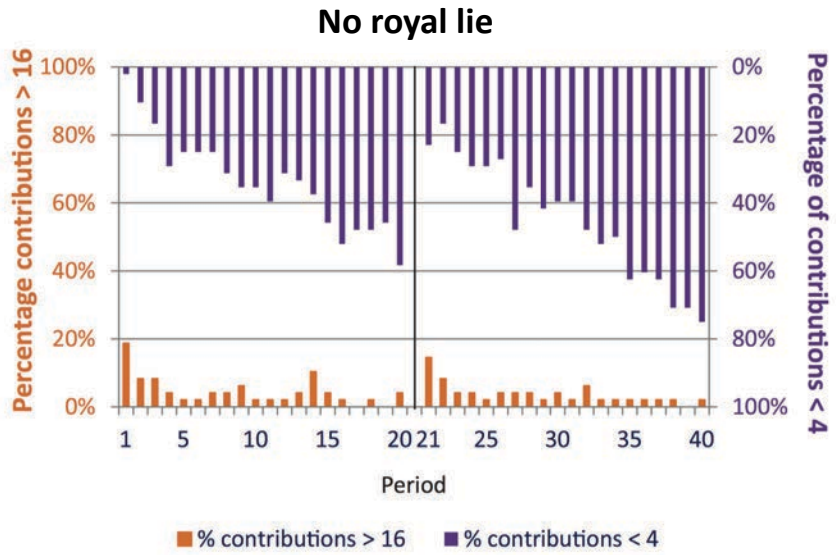
Between phases: Summary feedback about all periods and about whether information feedback deviated or not. Information that phase 2 will be in similar feedback condition than phase 1.

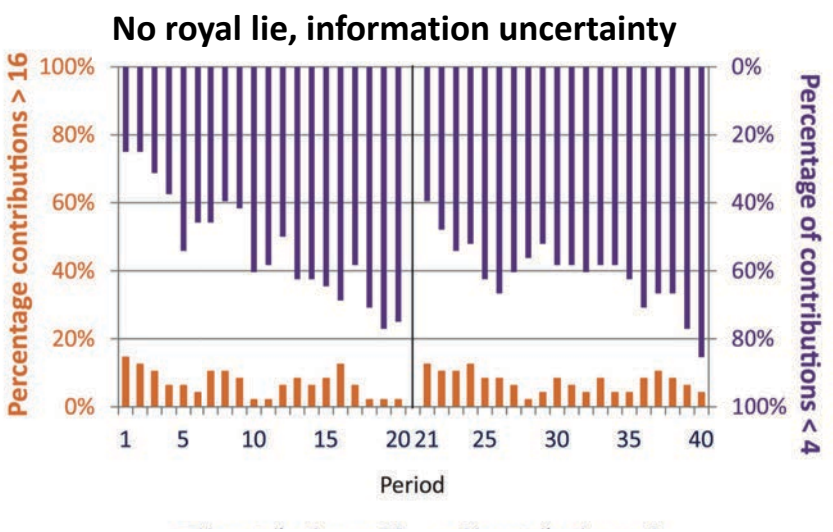
Fig. 4. Development of the percentage of low (<4) and high (>16) contributions.

The royal lie

Greek philosopher Plato “[. . .] if anyone at all is to have the privilege of lying the rulers of the State [. . .] may be allowed to lie for the public good” (The Royal Lie).

Hoffmann, Lauer & Rockenbach (2013, JEBO) experimentally show that manipulated feedback of public good provisions can maintain cooperation, in particular, when nobody receives the information of being exploited (i.e. being the sucker)





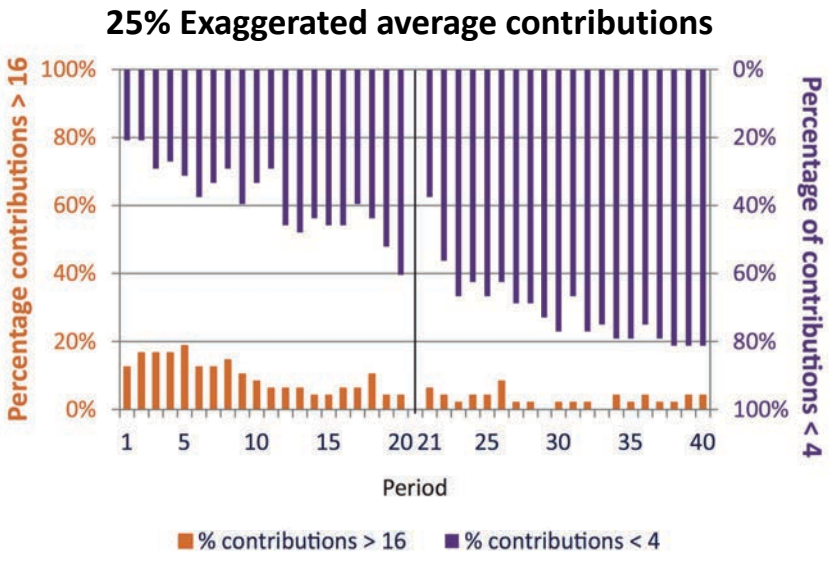
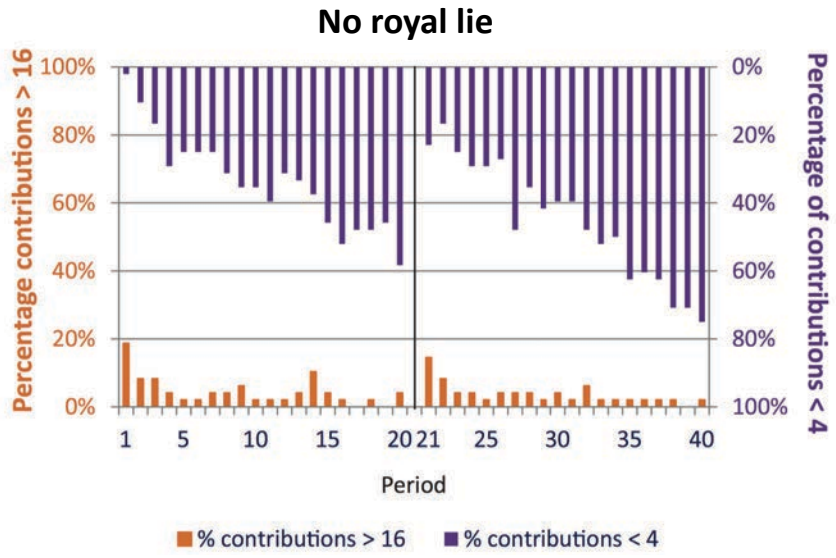


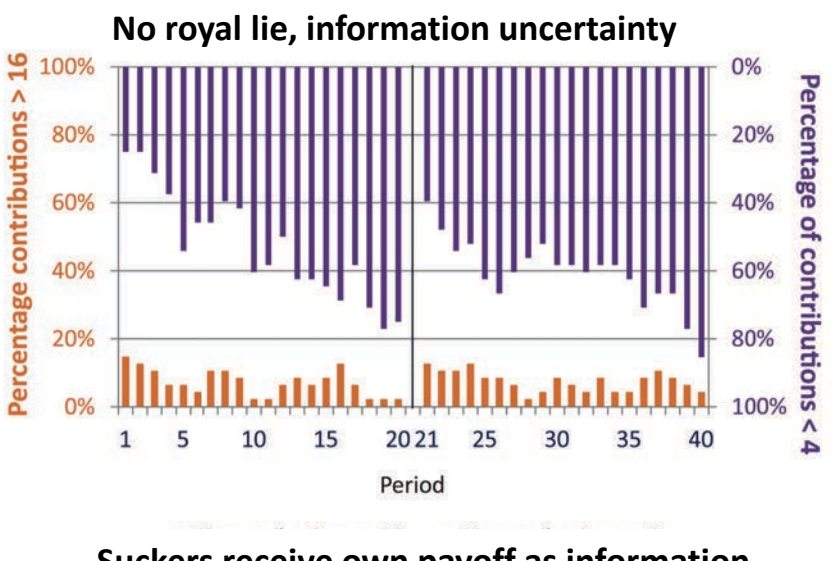
Fig. 4. Development of the percentage of low (<4) and high (>16) contributions.

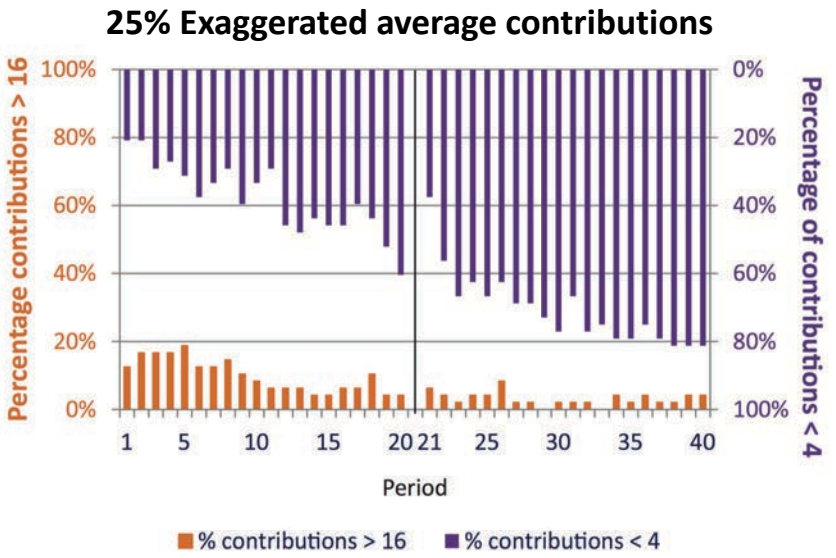
The royal lie

Greek philosopher Plato "[. . .] if anyone at all is to have the privilege of lying the rulers of the State [. . .] may be allowed to lie for the public good" (The Royal Lie).

Hoffmann, Lauer & Rockenbach (2013, JEBO) experimentally show that manipulated feedback of public good provisions can maintain cooperation, in particular, when nobody receives the information of being exploited (i.e. being the sucker)







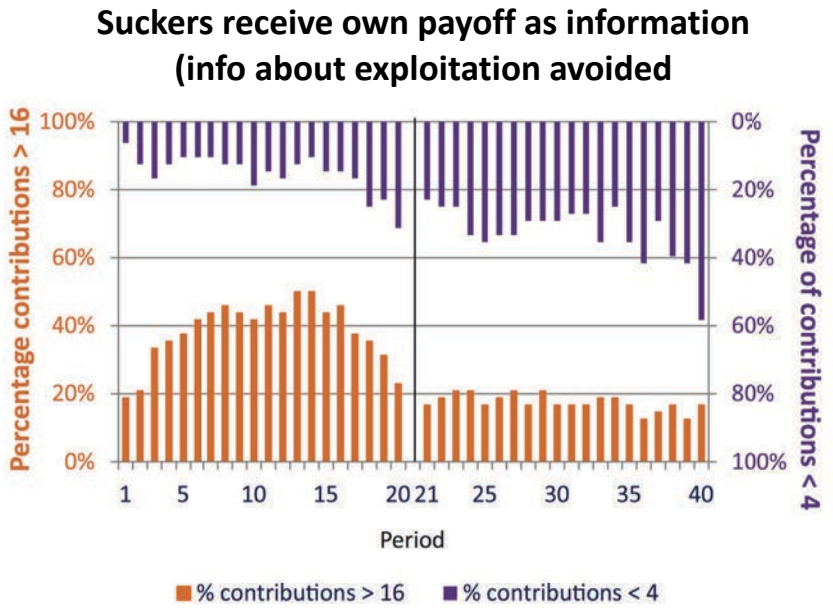


Fig. 4. Development of the percentage of low (<4) and high (>16) contributions.

The «first» experimental test of the ignorance hypothesis

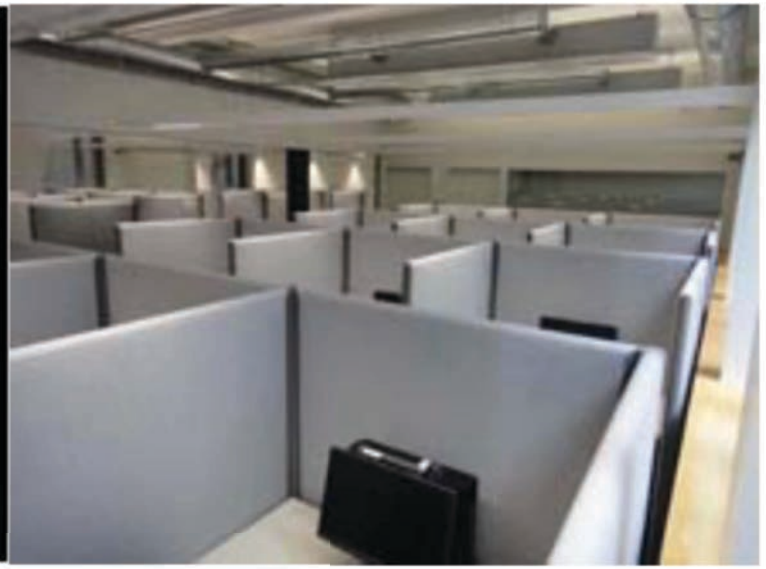
- **Counter-intuitive collective phenomenon**
  If all norm violations were detected (tax evation, fare-dogding, corruption, moonlighting, adultery, plagiarism etc.), norm violations would spread and systems of norms would collapse
- **Hypothesis**
  “Veil of ignorance” about norm violations prevents their spread (Popitz, 1968)
- **Assumption**
  People underestimate extent of norm violations
- **Research strategy**
  Isolation of underlying causal mechanisms

Diekmann, Przepiorka, Rauhut, Rationality & Society 2015

The «first» experimental test of the ignorance hypothesis

Dice experiments and the erosion of the honesty norm

Diekmann, Przepiorka, Rauhut, Rationality & Society 2015



- die cast of isolated and unobserved participants

die cast	1	2	3	4	5	6
CHF	1	2	3	4	5	0

Experimental design

- Norm: "Do not tell a lie!"
- Design based on Fischbacher & Heusi, 2008
- Isolated and unobserved subjects throw a die
- Cash according to their reported number in Swiss Francs
- No earnings for number six
- Tension between lying and adherence to the honesty norm for numbers other than five
- 30 sessions, 14-16 subjects in each, 466 subjects in total

Table: Experimental treatments

info 0: control group	◯		◯
info 1: "external" distribution of 1st throws	◯	$\times_1$	◯
info 2: "internal" distribution of 1st throws	◯	$\times_2$	◯

Diekmann, Przepiorka, Rauhut, Rationality & Society 2015

Instruction for 1st die cast

Please cast the die now!

To check whether the die is working correctly, you may cast it several times. But **only the first cast counts**.

Enter the number on the die of your first cast and your payoff according to the table below in the corresponding fields.

Number on the die: 3

Payoff in CHF: 3

Number on the die:	1	2	3	4	5	6
Payoff in CHF:	1.00	2.00	3.00	4.00	5.00	0.00

OK

Individual classroom task

Imagine a group of 60 students playing the dice game for one round. What is your best guess for the distribution of claimed payoffs?

Please write down on a sheet of paper the number of estimated observations for each of the six cells, such that they add up to 60. Later we collect all estimates on the flip chart.

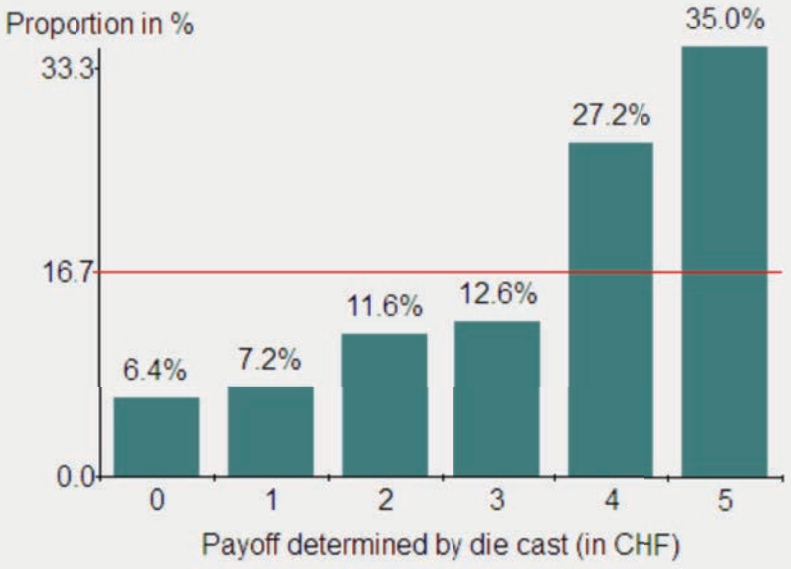
Number on the die:	1	2	3	4	5	6
Payoff in CHF:	1.00	2.00	3.00	4.00	5.00	0.00

Treatment info 1 ("external" distribution)

This graph shows the generated distribution of payoffs **by 389 students** from ETH and University of Zurich who participated in the same experiment.

The **red line** denotes the average proportion of payoffs which would have resulted from a **large number of random die casts**.

Subsequently you receive the opportunity to cast the die again. The payoff determined by your next cast will be added to your previous payoff.



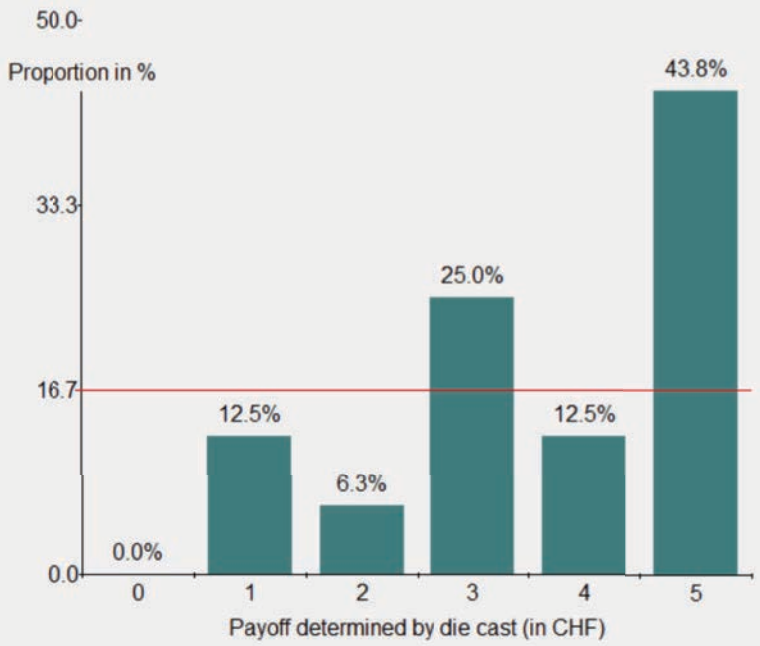
OK

Treatment info 2 ("internal" distribution)

This graph shows the generated distribution of payoffs **by the 16 participants** in this experimental session.

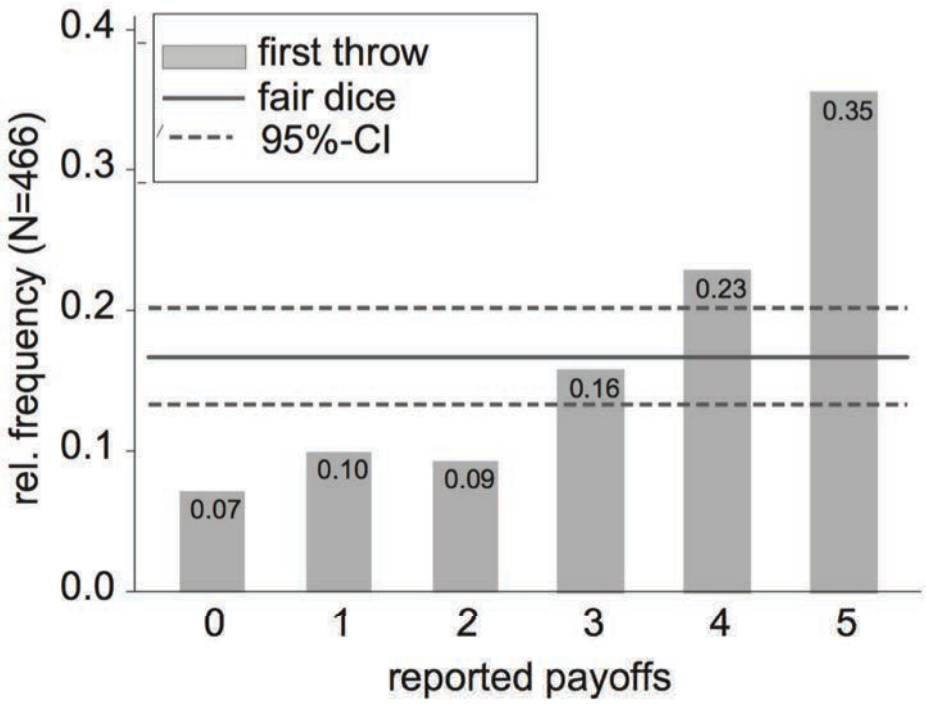
The **red line** denotes the average proportion of payoffs which would have resulted from a **large number of random die casts**.

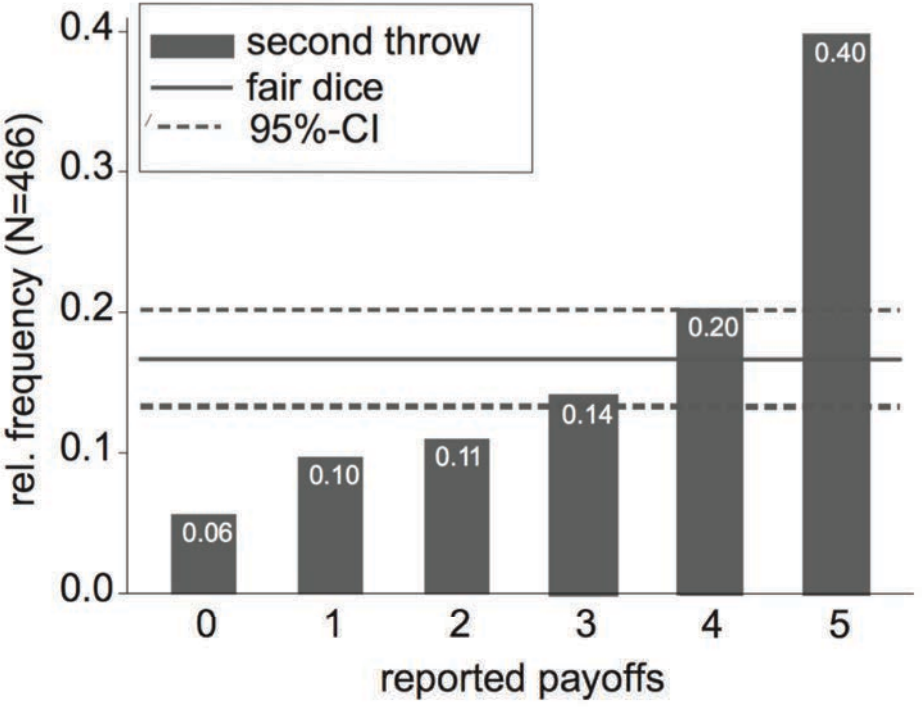
Subsequently you receive the opportunity to cast the die again. The payoff determined by your next cast will be added to your previous payoff.



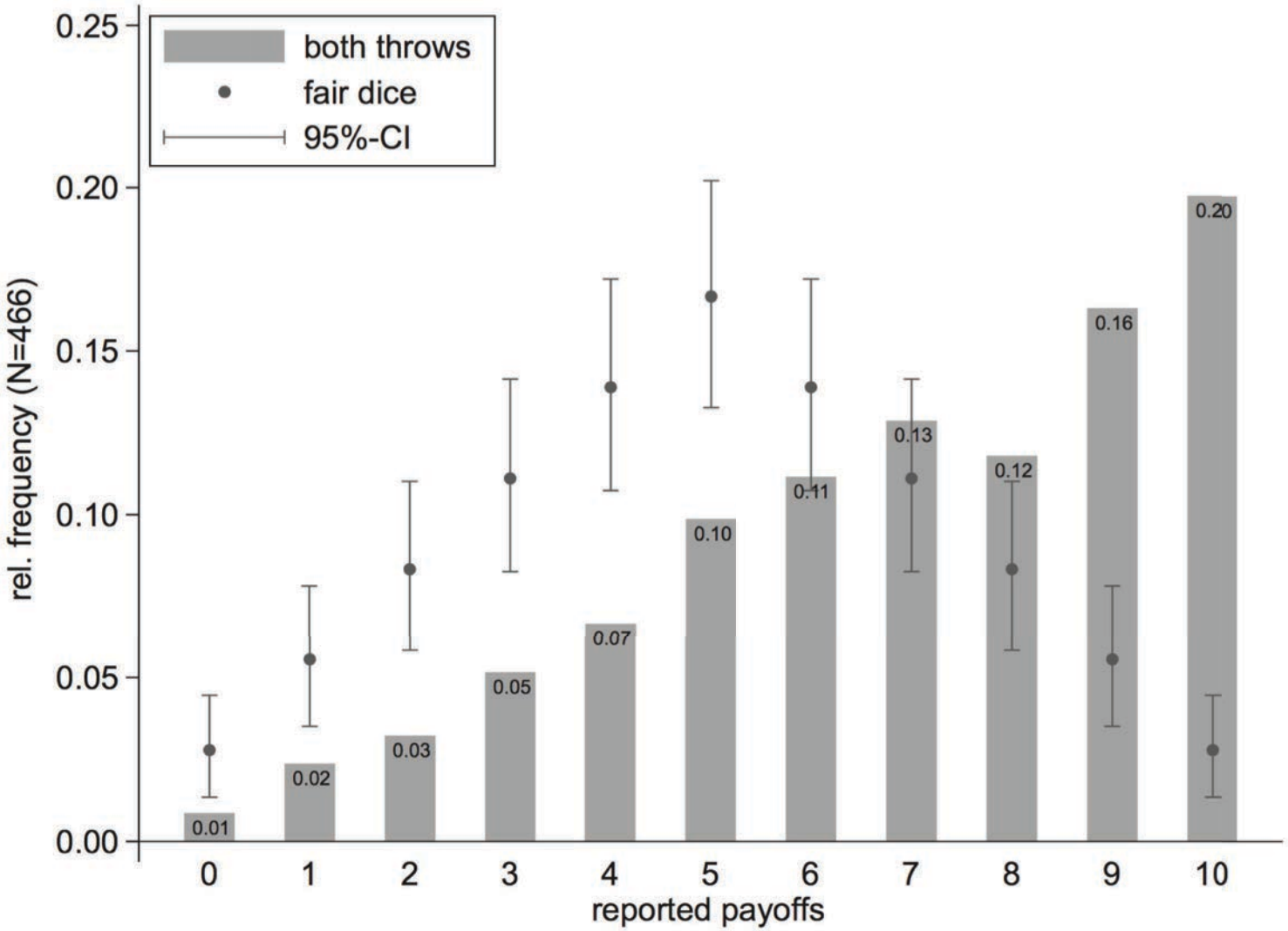
OK

Distribution of reported 1st and 2nd throws
(averaged over all treatments)





Distribution of reported cumulated payoffs (averaged over all treatments)



Information on lying triggers subsequent lying

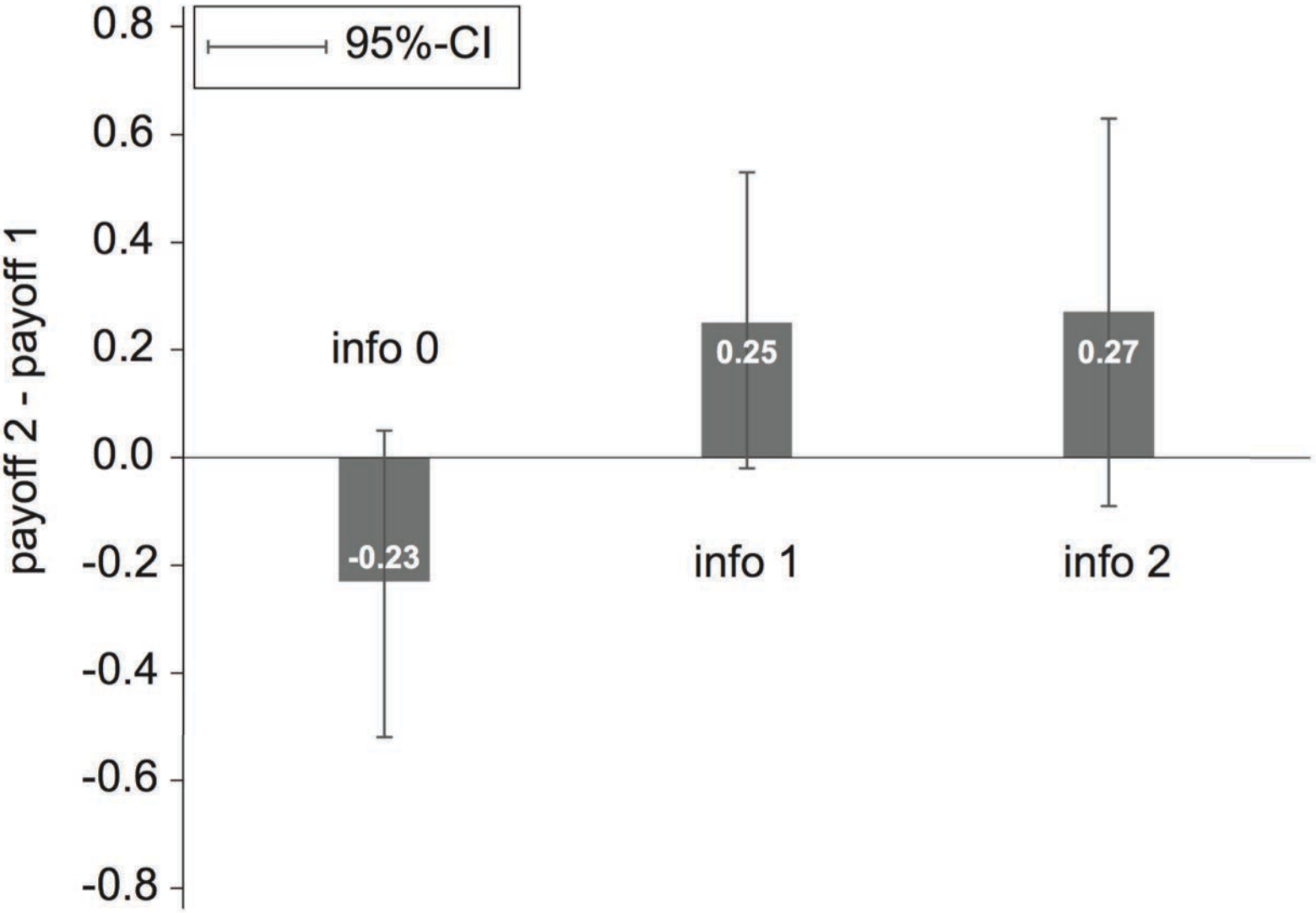


Figure: Difference between 1st and 2nd reported throws

Information on lying triggers subsequent lying

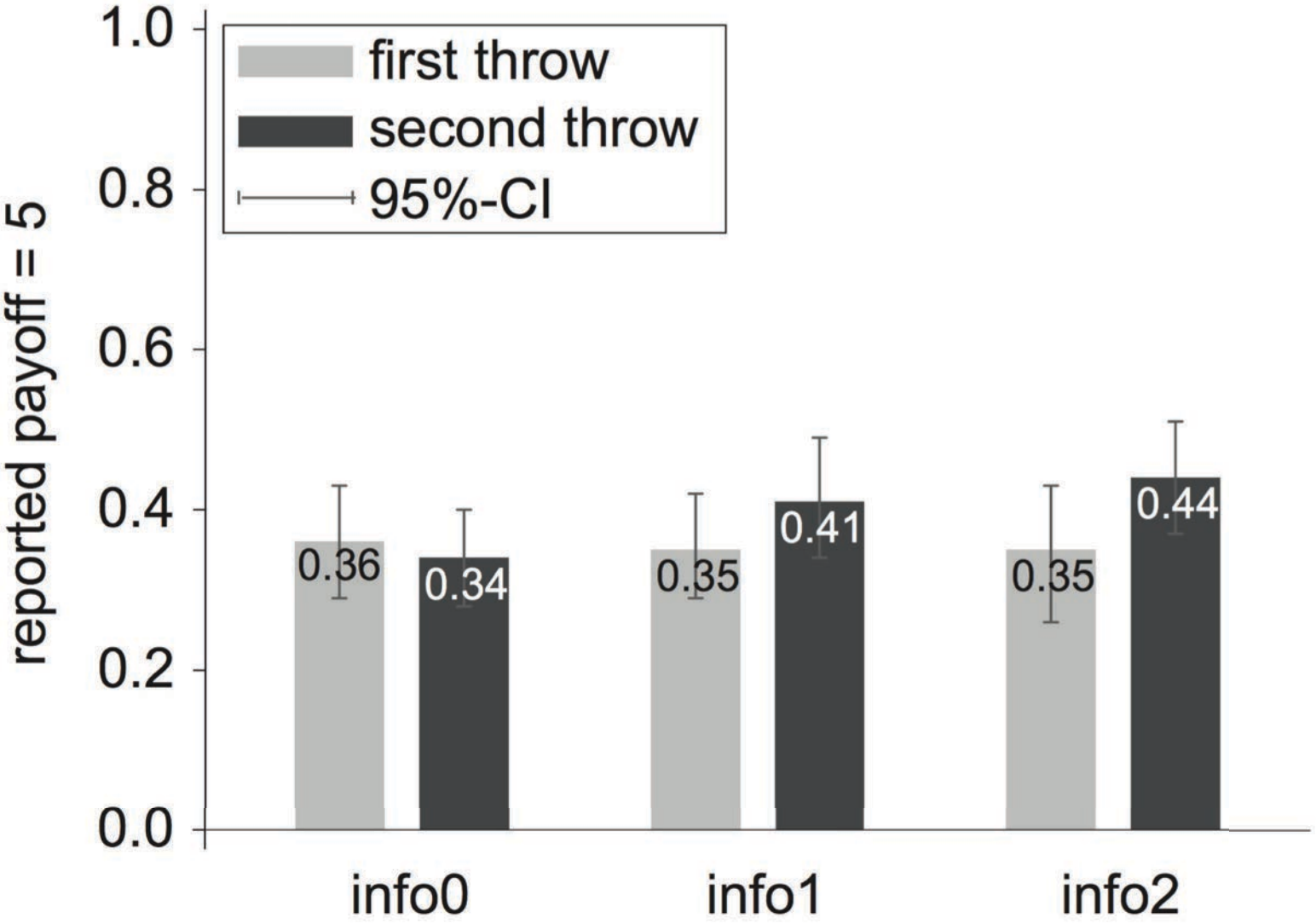


Figure: Fraction of reported "five" in 1st and 2nd throw by treatment

Information on lying triggers subsequent lying

	OLS 1 (Ausz. 2 – Ausz. 1)		OLS 2 (Ausz. 2 – Ausz. 1)		Logit (Ausz. = 5)	
	Koef.	SF	Koef.	SF	Koef.	SF
first cast					(ref.)	
info 0 (control group)	(ref.)		-0.234	0.138	-0.054	0.167
info1 (external)	0.487*	0.195	0.253	0.135	0.243	0.206
Info 2 (internal)	0.501*	0.189	0.267	0.177	0.360*	0.182
intercept	–0.234	0.138			–0.601*	0.105
N_1 (subjects)	466		466		932	
N_2 (sessions)	30		30		30	

Clustered standard errors (* $p < 0.05$). Coefficients in the logit model measure the change in the likelihood to report a five in the second throw.

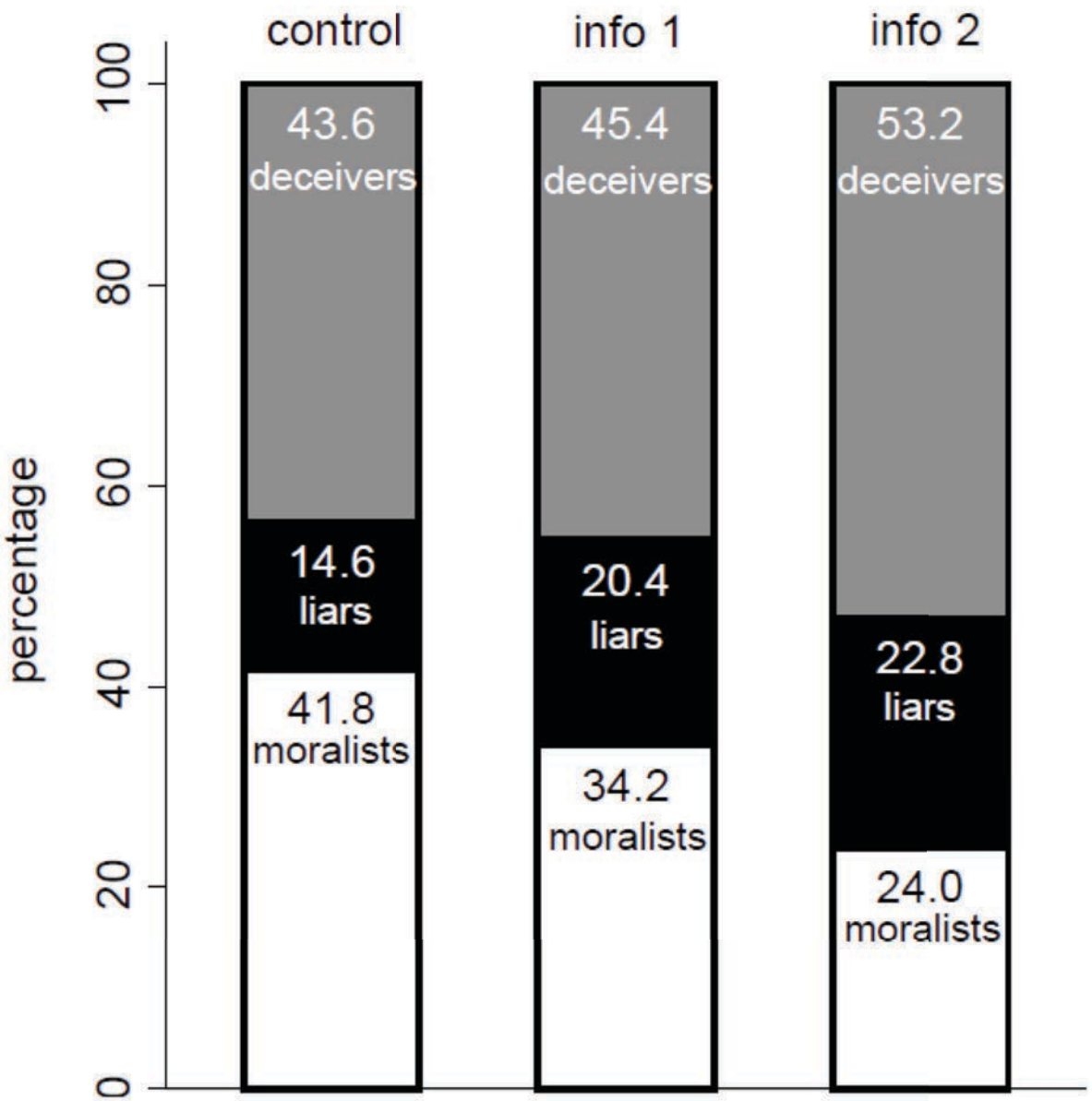
Model 1, joint test, ANOVA, p=0.015; Model 2, joint test, ANOVA, p=0.029

Model 1, joint test, ANOVA, p=0.015
Model 2, joint test, ANOVA, p=0.029

Classification of types (Fischbacher & Heusi, 2008)

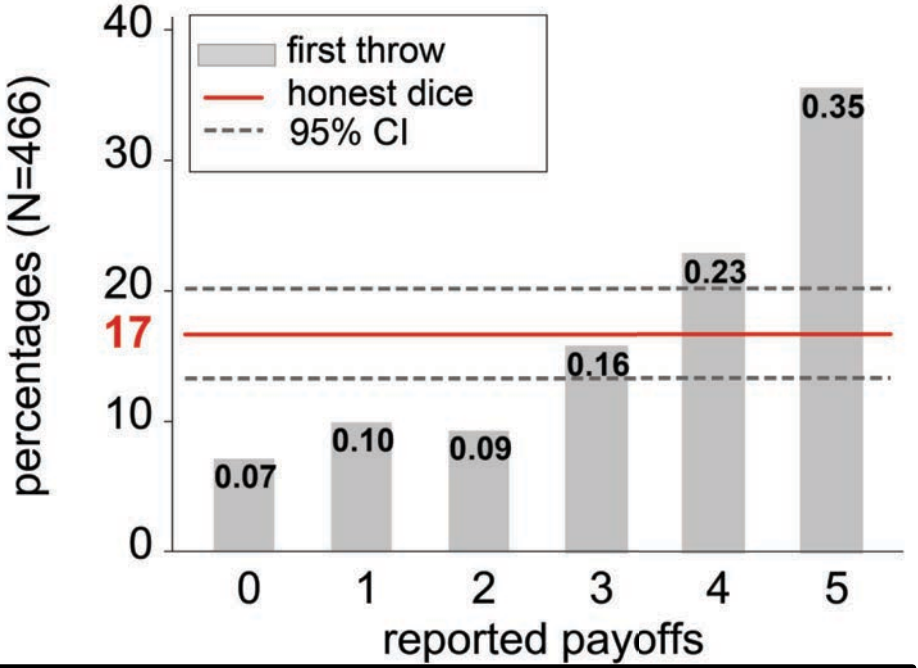
1. moralists
   - report truth regardless of entitlement
   - proportion zero payoffs × six
2. complete liars
   - report highest payoff regardless of entitlement
     - ★ $0 \rightarrow 5$
     - ★ $1 \rightarrow 5$
     - ★ $2 \rightarrow 5$
     - ★ $3 \rightarrow 5$
     - ★ $4 \rightarrow 5$
   - difference between expected percentage of five's (1/6) and empirically reported proportion of five's ("cf")
   - adjustment for liars who actually threw a five (multiple of 6/5)
   - (cf - 1/6) * 6/5
3. deceivers
   - increase partially their payoffs to disguise their lies
     - ★ $0 \rightarrow 1, 0 \rightarrow 2, 0 \rightarrow 3, 0 \rightarrow 4$
     - ★ $1 \rightarrow 2, 1 \rightarrow 3, 1 \rightarrow 4$
     - ★ $2 \rightarrow 3, 2 \rightarrow 4$
     - ★ $3 \rightarrow 4$
   - assumption: remaining population (100% - moralists - liars)

Fig 6. Proportions of different types in second throws by experimental conditions



The «first» experimental test

Diekmann, Przepiorka, Rauhut,
Rationality and Society, 2015



- One die cast before and one after information feedback
- Comparison of information treatments about others' lying in large «stranger» group ($n > 350$) and small own groups ($n \geq 14$) with control condition without information feedback
- (Modest) confirmation of ignorance hypothesis: More lying after information about others' lies compared to control condition

What about those who overestimate lying?

Rauhut, PLoS One, 2013

- **Is the dynamics inverted (less transgressions instead of more) if informed about true rate?**
  - **"Underestimators" (standard assumption)**

    perceive public occurrences of others' norm violations as relatively frequent or strong, increase their subjective estimates about the complete extent of norm violations and perform subsequently more own norm violations
  - **«Overestimators» (extended assumption)**

    perceive public occurrences of others' norm violations as relatively rare or mild, decrease their subjective estimates about the complete extent of norm violations and perform subsequently less own norm violations

- **Interaction effect between beliefs and direction of normative dynamics**

– information about norm violations triggers increasing norm violations for underestimators, and decreasing norm violations for overestimators

Experimental design

Casted number	6	1	2	3	4	5
Payment in CHF	0.00	1.00	2.00	3.00	4.00	5.00

This is the start of the main study. From now on, your entries are payment relevant. One of your die casts will be randomly selected for payments in cash.

Please cast 12 times your die and fill in your scored points into the following table.

	Augenzahl	6	1	2	3	4	5
payment →	Auszahlung	0 CHF	1 CHF	2 CHF	3 CHF	4 CHF	5 CHF
cast 1 →	Wurf 1	●	○	○	○	○	○
cast 2 →	Wurf 2	●	○	○	○	○	○
cast 3 →	Wurf 3	○	○	○	○	●	○
cast 4 →	Wurf 4	○	○	○	●	○	○
cast 5 →	Wurf 5	○	○	○	○	○	●
cast 6 →	Wurf 6	○	●	○	○	○	○
cast 7 →	Wurf 7	○	○	●	○	○	○
cast 8 →	Wurf 8	●	○	○	○	○	○
cast 9 →	Wurf 9	○	○	○	●	○	○
cast 10 →	Wurf 10	○	○	○	○	○	●
cast 11 →	Wurf 11	○	○	○	○	○	●
cast 12 →	Wurf 12	○	○	○	○	○	○

continue → Weiter

Payments

- 1 cast randomly paid out per round
- 4 payment rounds with 12 casts each

Sample

- 24 groups, each of which 10 subjects (N=240)
- Students, ETH & University of Zurich

Treatments

- control base
- control belief
- info

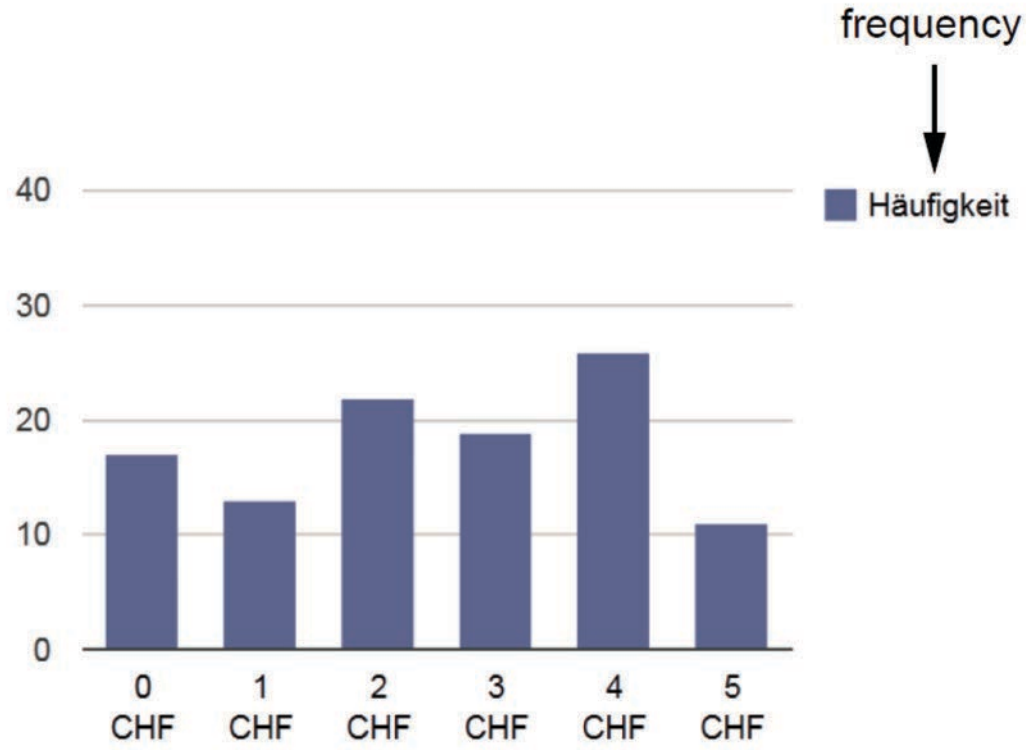
Why multiple dice casts?

- elicitation of beliefs in each round of each session
- dice reports of only 9 other group members robust elicitation of meaningful beliefs: 12x9 = 108 dice casts each session

Belief elicitation

reported cast → | frequency →

Auszahlung	Häufigkeit
0 CHF	17
1 CHF	13
2 CHF	22
3 CHF	19
4 CHF	26
5 CHF	11
total number →	Gesamtzahl: 108

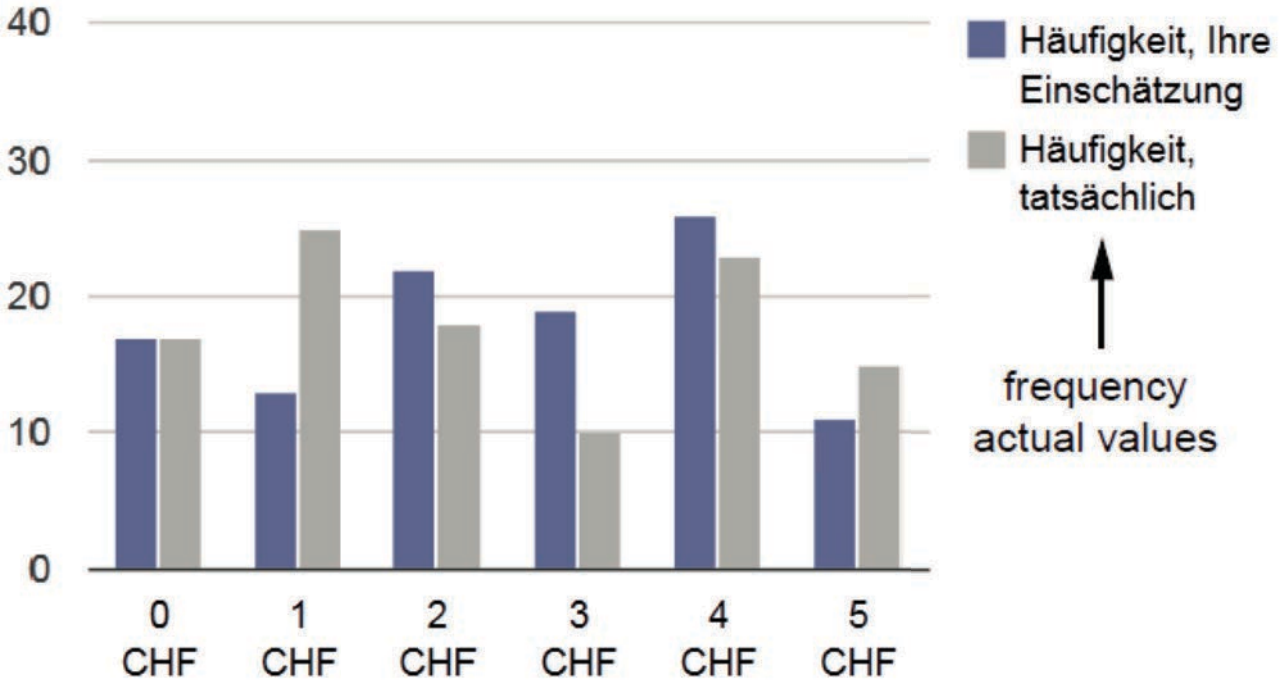


Weiter ← continue

Difference belief and real frequency of reported payoff	0	1	2	3	≥ 4
CHF	0.80	0.75	0.60	0.35	0

Information feedback

The following figure shows your estimates in comparison to the actual values.



Für Ihre Einschätzung erhalten Sie 1.15 CHF. ← Your receive 1.15 CHF for your estimates.

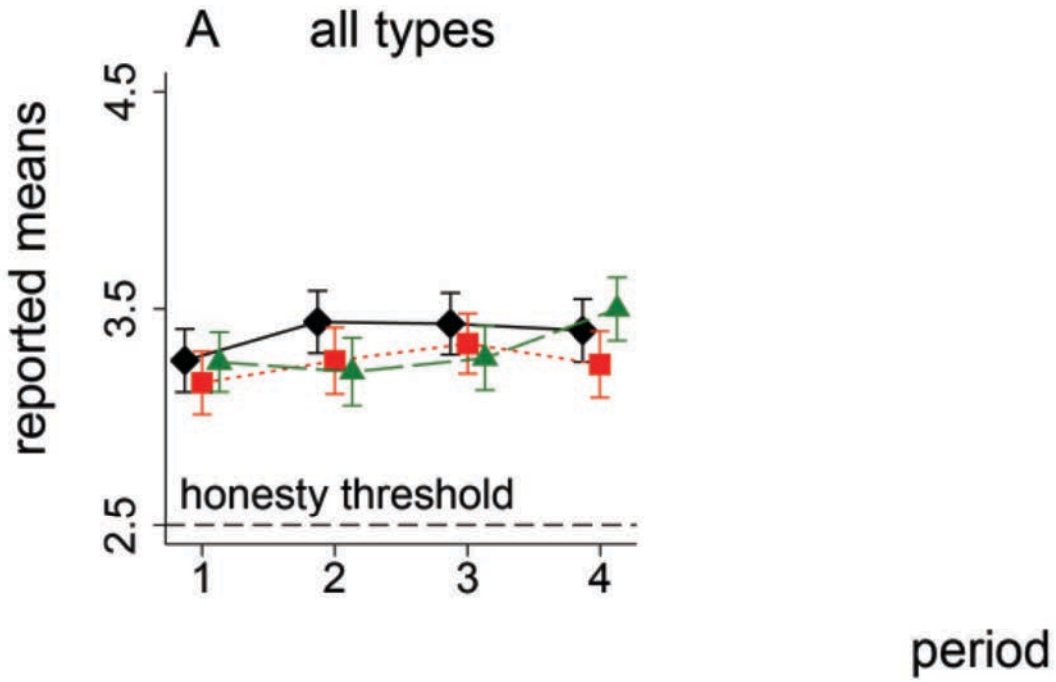
Weiter ← continue

Random assignment to three treatments
(within each session)

treatment	belief elicitation	information feedback
info	☒	☒
control belief	☒	
control base		

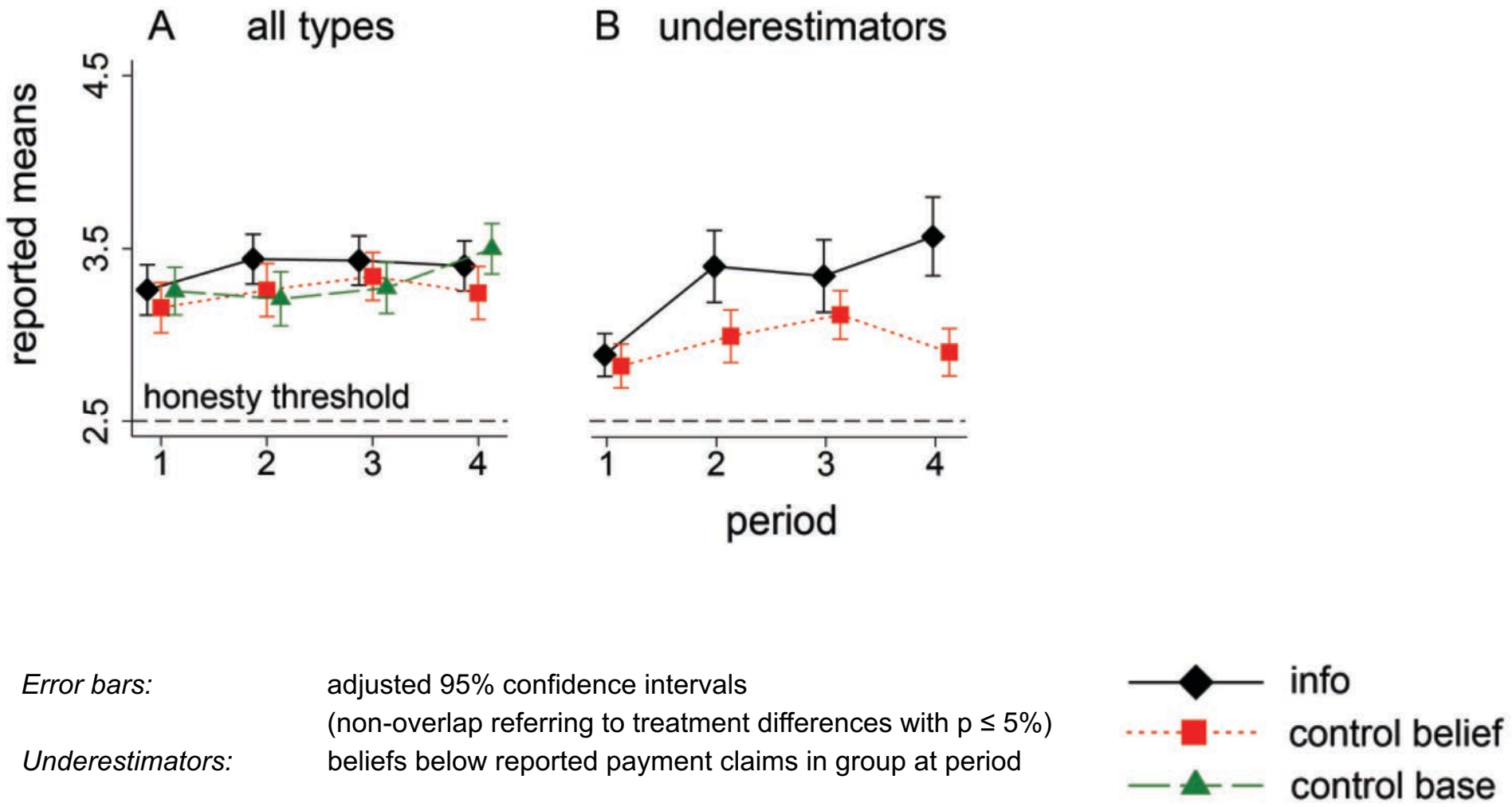
Trend of reported payment claims in means

Trend of reported payment claims in means



Error bars: adjusted 95% confidence intervals
(non-overlap referring to treatment differences with p ≤ 5%)

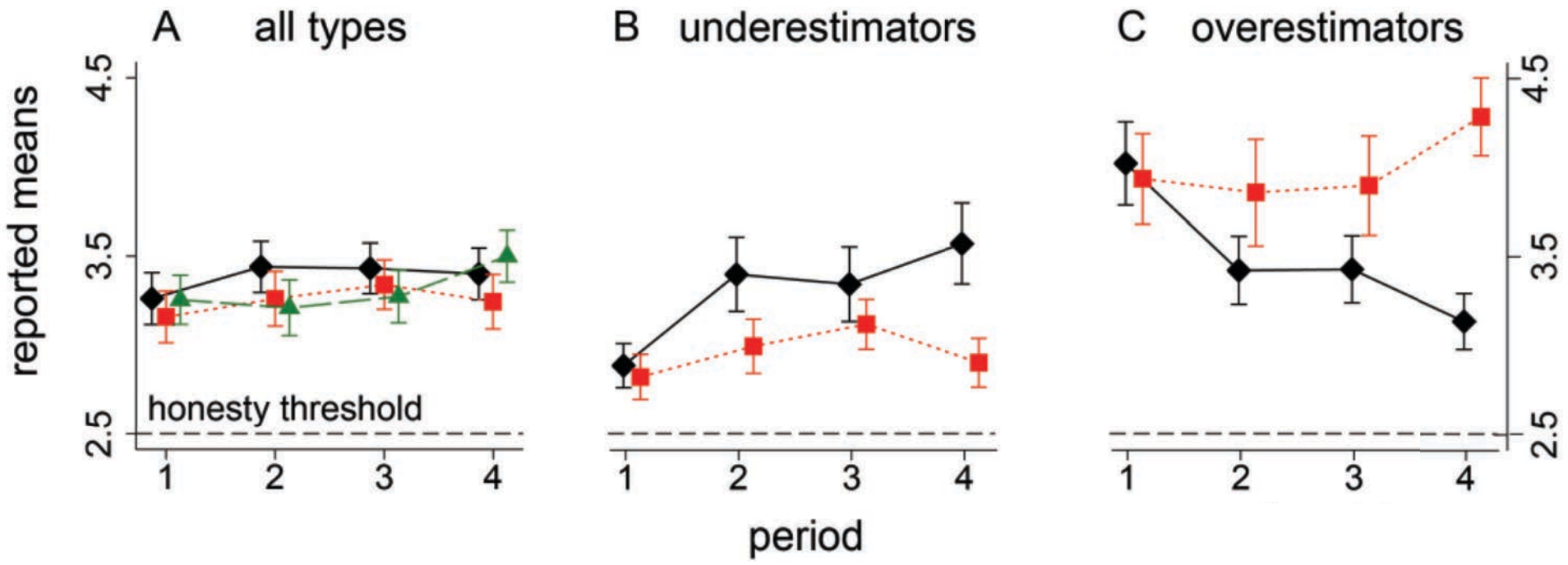
Trend of reported payment claims in means



Error bars: adjusted 95% confidence intervals
(non-overlap referring to treatment differences with p ≤ 5%)

Underestimators: beliefs below reported payment claims in group at period

Trend of reported payment claims in means

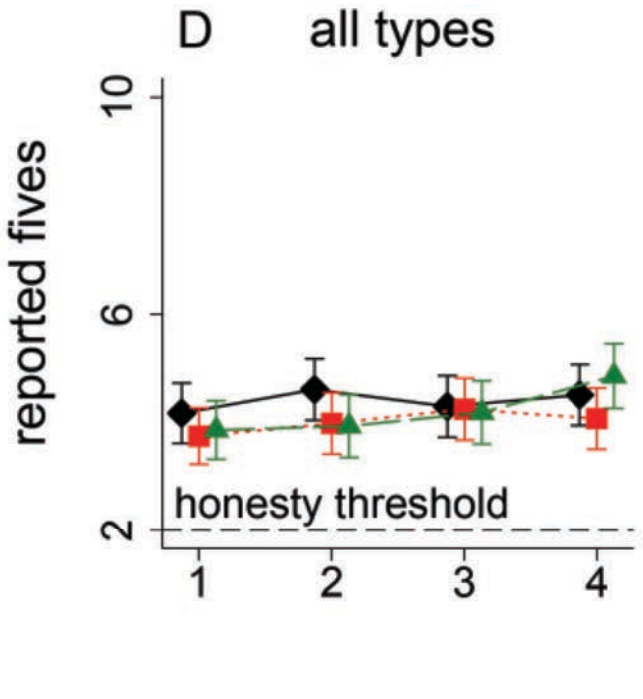


Error bars: adjusted 95% confidence intervals
(non-overlap referring to treatment differences with p ≤ 5%)

Underestimators: beliefs below reported payment claims in group at period

Overestimators: beliefs above reported payment claims in group at period

Trend of reported payment claims in fives

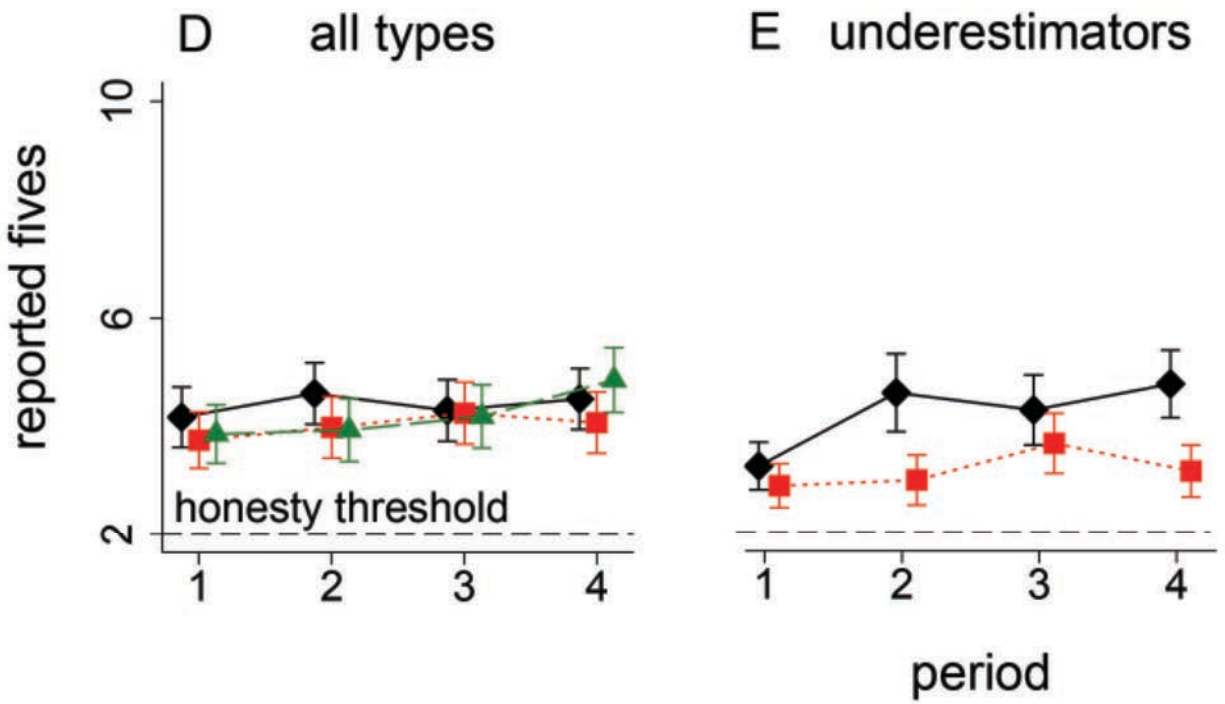


Error bars: adjusted 95% confidence intervals
(non-overlap referring to treatment differences with p ≤ 5%)

Underestimators: beliefs below reported payment claims in group at period

Overestimators: beliefs above reported payment claims in group at period

Trend of reported payment claims in fives

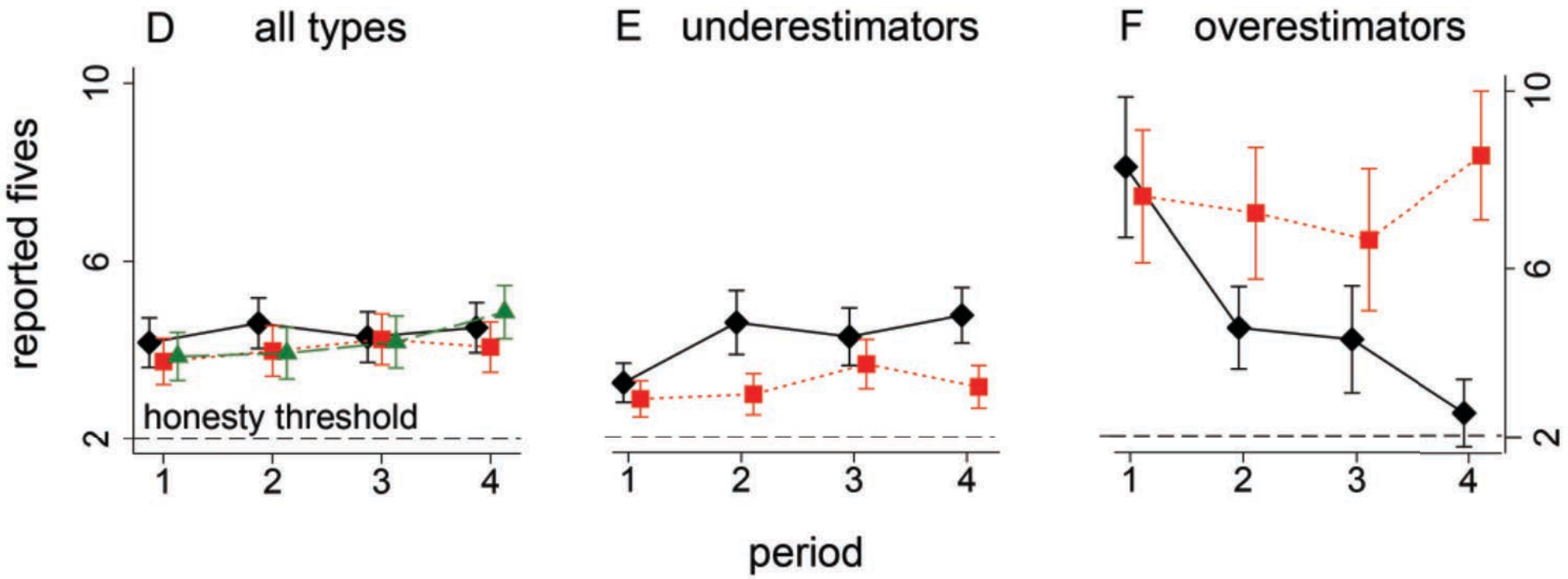


Error bars: adjusted 95% confidence intervals
(non-overlap referring to treatment differences with p ≤ 5%)

Underestimators: beliefs below reported payment claims in group at period

Overestimators: beliefs above reported payment claims in group at period

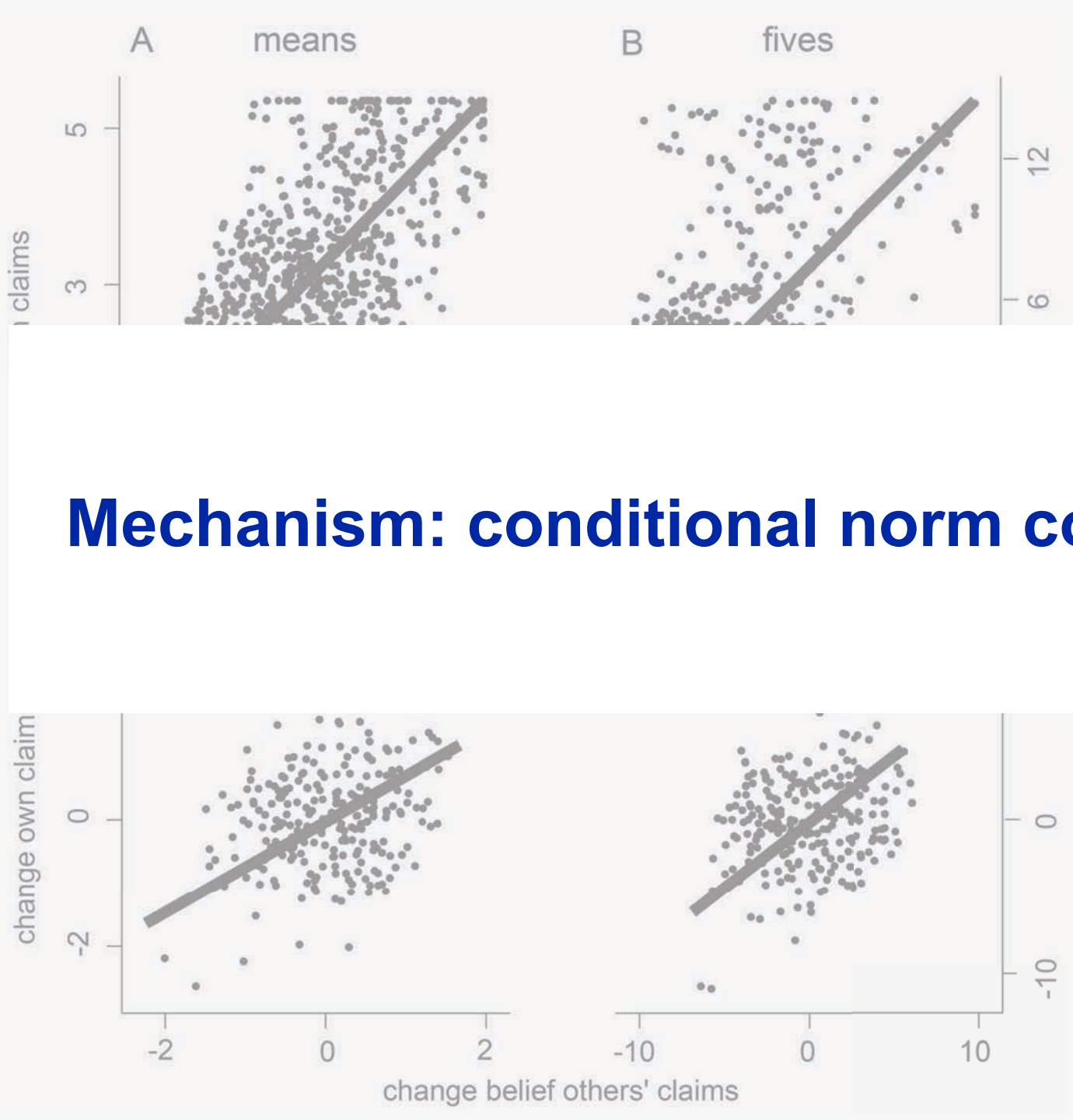
Trend of reported payment claims in fives



Error bars: adjusted 95% confidence intervals
(non-overlap referring to treatment differences with p ≤ 5%)

Underestimators: beliefs below reported payment claims in group at period

Overestimators: beliefs above reported payment claims in group at period



Mechanism: conditional norm compliance

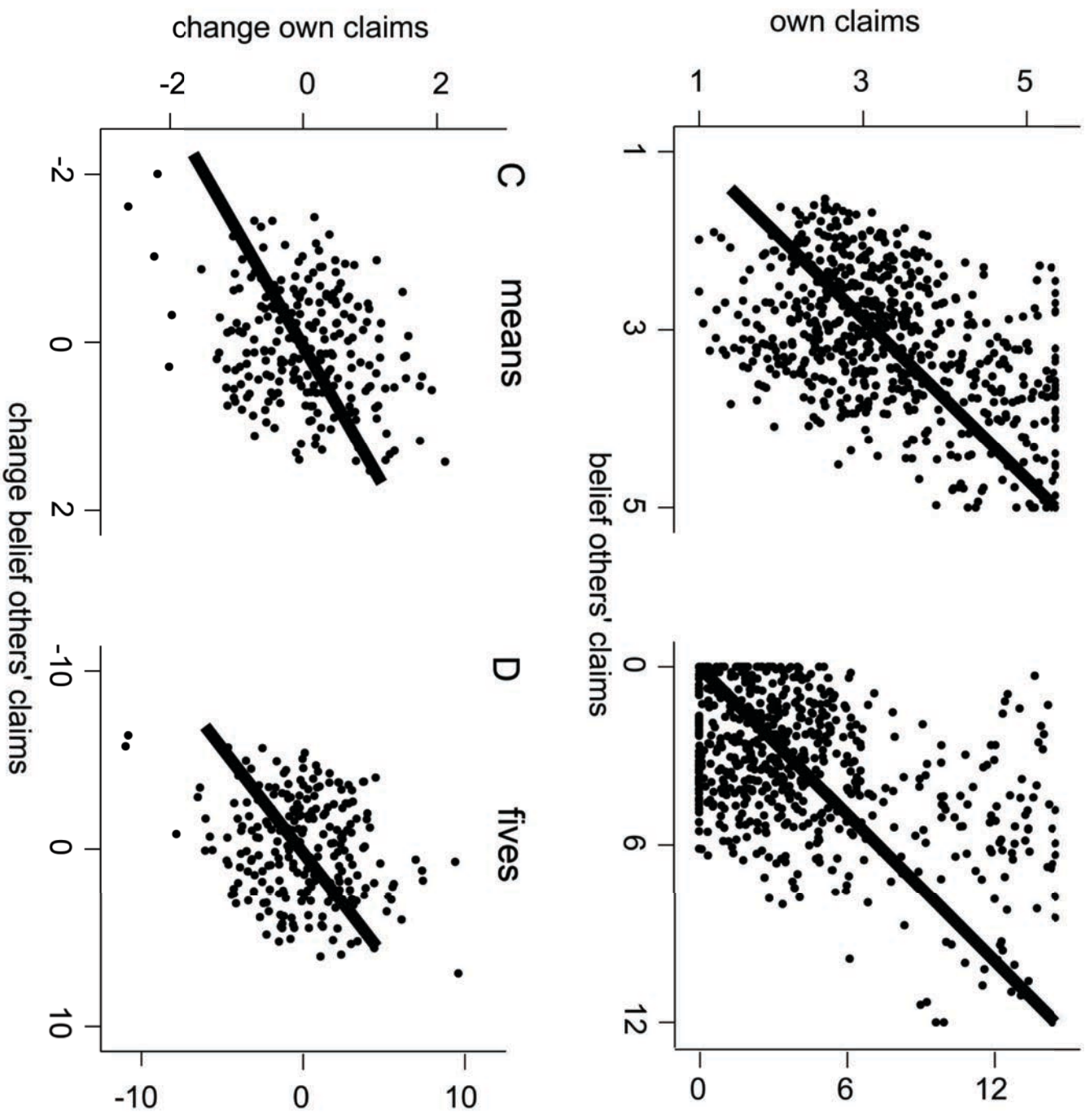


Figure 2: Scatterplots and linear regression lines on conditional honesty. The upper two panels show beliefs' about other claims on own claims in the same period for reported means (panel A) and reported fives (panel B). The upper panels consist of info and control belief treatments and periods are pooled for t=1,2,3,4. The lower two panels show the correlation of the change in beliefs of others' claims from period t to period t+1 on the change in own claims from period t to period t+1 in reported means (panel C) and reported fives (panel D). The lower panel consist of the info treatment and changes refer to pooled periods 1–2, 2–3, and 3–4.

Conditional liars

- Estimation of percentage of «complete» liars:
  - expected proportion of the highest payoff five of a fair die (1/6)
  - compare to reported proportion of fives π
  - adjust for liars who threw five, but would have lied for lower numbers (i.e. multiply by 6/5).
  - proportion of liars $\lambda = (\pi - 1/6) \cdot 6/5$

Lying can be more than halved or more than doubled depending on subjective beliefs and information feedback

- Underestimators: Twice as much liars in info (25.6 %) than control (12.7%)[1]
- Overestimators: Less than half liars in info (21.8%) than control (56.3%)[1]

[1] *percentages refer to periods 2-4 after information feedback in info and control belief*

Discussion: Implications for actor models

- Learning and macro dynamics unexplainable by homo oeconomicus
  - Rational learning of punishment severity and probability eliminated by design
  - Homo oeconomicus no dynamics: always maximum claim

- «Modern homo sociologicus»
  - Actors follow norms conditional on norm compliance of others
  - Actors are both, self-regarding and other-regarding
  - Evidence compatible with «Beliefs, Preferences, Restrictions» (Bowles/Gintis)

- «Self-serving bias»?
  - Only self-serving learning in gift exchange games:
    information about others' violations of reciprocity norms has only effects for normative decay (*Thöni & Gächter, 2012)*
  - Dice experiments show learning in both directions; constructive and destructive dynamics

Group work

- **Group 1: Discuss and present examples of the ignorance hypothesis with:**

  Counter-intuitive collective phenomenon
  If all norm violations were detected, norm violations would spread

  Hypothesis
  "Veil of ignorance" about norm violations prevents their spread

  Micro- level assumption
  People underestimate extent of norm violations

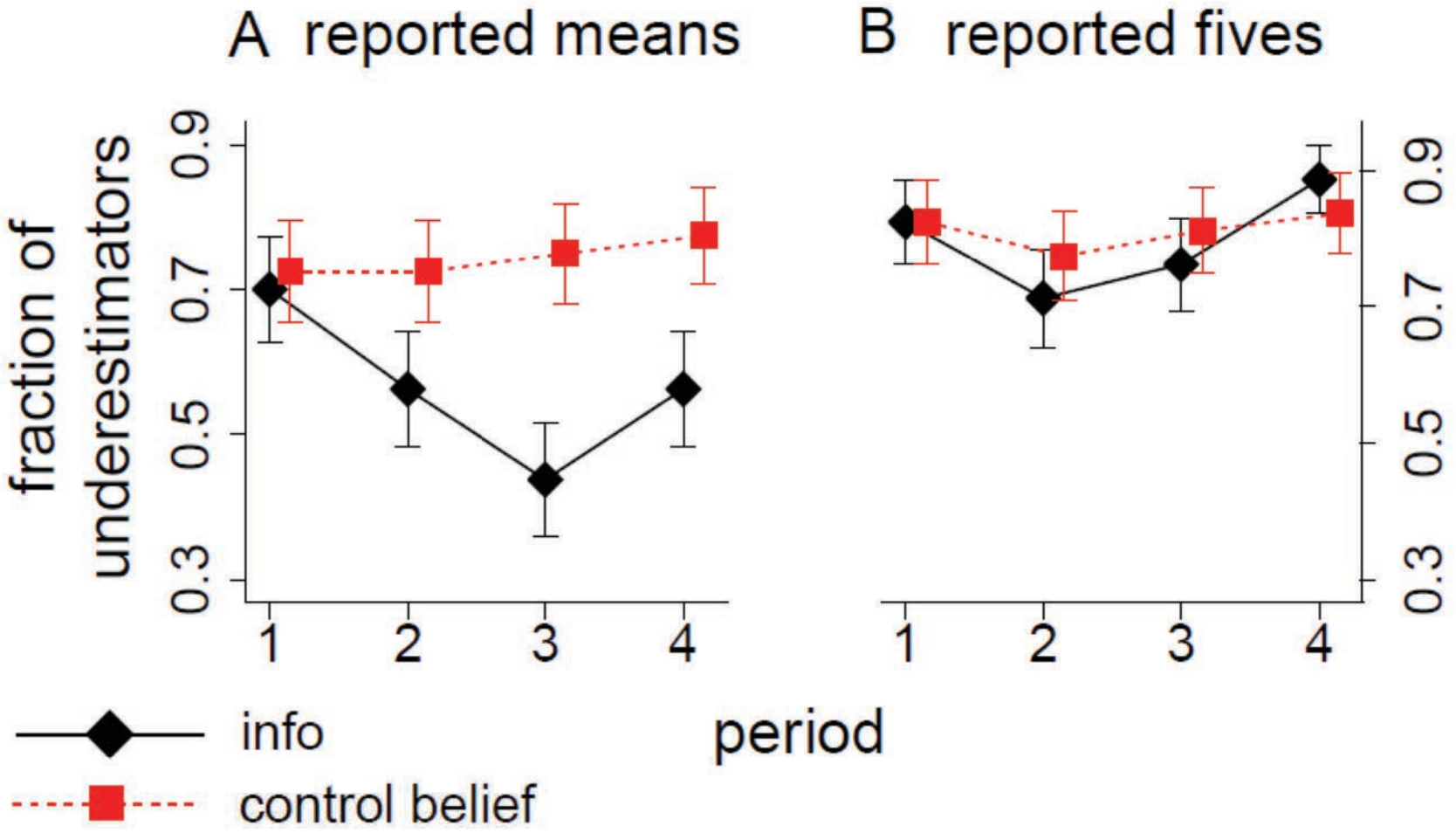
- **Group 2: Discuss and present a design for field experiments of cross-norm inhibition effects (broken windows) with:**

  Several examples of 2 norms: the displayed norm violations of norm 1 and the cross-norm inhibition effect on norm 2

Appendix

Fraction of under- and overestimators in PLoS Study

Fig S6. Group sizes of under- and overestimators over periods. Panel A displays the fraction of underestimators of reported means and panel B, of reported fives. Error bars show adjusted 95% confidence intervals such that non–overlapping intervals refer to treatment differences with $p \leq 5\%$ (see SM for calculations of adjustments). Underestimators hold beliefs below reported payment claims in their group at respective periods.



Significance and effect sizes for belief dynamics

Linear regression models of treatment differences

	(A) means	(B) fives
info	-0.715 ***	-3.454 **
	(-3.72)	(-3.30)
underestimator types	-1.114 ***	-4.362 ***
	(-6.39)	(-4.51)
info × underestimator types	1.156 ***	4.741 ***
	(5.19)	(4.31)
intercept	4.118 ***	7.630 ***
	(25.47)	(8.18)
N	480	480

Model A shows differences in claimed mean payments and model B differences in claimed number of fives with respect to under- and overestimators and their treatment interactions. One case refers to the reported mean (model A) or reported number of fives (model B) over the sequence of twelve dice casts per period per subject (yielding a total of N=480 cases for each model). Only periods 2, 3 and 4 are used, because these are the periods after information feedback in the info treatment. Robust standard errors are used, which were clustered for subjects. T statistics are reported in parenthesis, stars denote statistical significance with $*$ $p < 0.05$, $**$ $p < 0.01$, $***$ $p < 0.001$.