Discussion Papers of the Max Planck Institute for Research on Collective Goods Bonn 2010/06

Turning the Lab into Jeremy Bentham's Panopticon

A Lab Experiment on the Transparency of Punishment

Christoph Engel

MAX PLANCK SOCIETY

Discussion Papers of the
Max Planck Institute
for Research on Collective Goods

Bonn 2010/06

Turning the Lab into Jeremy Bentham's Panopticon

A Lab Experiment on the Transparency of Punishment

Christoph Engel

January 2010

Revised version: June 2018

Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, D-53113 Bonn
http://www.coll.mpg.de

Turning the Lab into Jeremy Bentham's Panopticon

A Lab Experiment on the Transparency of Punishment

Christoph Engel[*]

Abstract

The most famous element in Bentham's theory of punishment, the Panopticon Prison, expresses his view of the two purposes of punishment, deterrence and special prevention. This paper investigates Bentham's intuition in a public goods lab experiment, by manipulating how much information on punishment experienced by others is available to would-be offenders. Compared with the tone that Jeremy Bentham set, the result is non-expected: If would-be offenders learn about contributions and punishment of others at the individual level, they contribute much less to the public project.

Keywords: Punishment, Deterrence, Special Prevention, Jeremy Bentham, Experiment, Public Good

JEL: C91, H41, K14, K42

* Prof. Dr. Christoph Engel, Max-Planck-Institute for Research on Collective Goods, Bonn. This paper has started from a joint project with Bernd Irlenbusch. Due to other obligations, he has withdrawn from the project. Research assistance by Karsten Lorenz and Lilia Zhurakhovska and helpful comments by Markus Englerth and Andreas Nicklisch are gratefully acknowledged.

I. Research Question

Not many scientific achievements are cast in stone. Jeremy Bentham's theory of punishment is among the exceptions. Inspired by his younger brother Samuel, a naval architect based in Russia in the 1780s, Jeremy Bentham translated his theory into a blueprint for the design of prisons (Semple 1993). Over time, all over the world some 20 "panopticon" prisons have been built[1]. The basic idea is simple. The circular construction places each and every inmate under permanent control by the supervisor, located at the centre. Every spectator cannot but realize that prisoners have fully lost their autonomy.

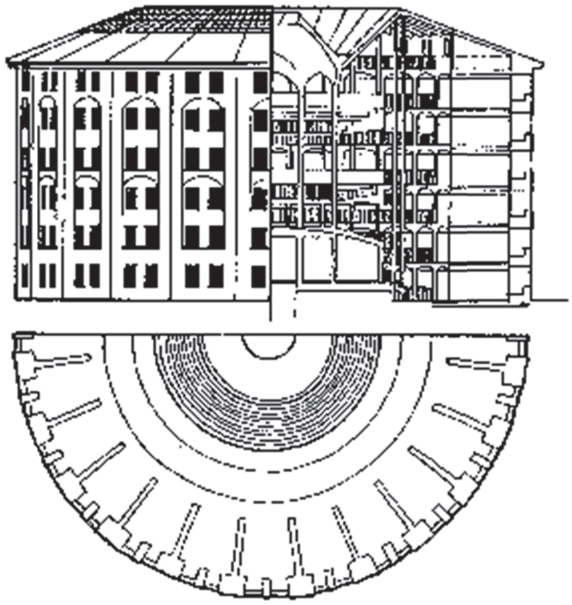


Figure 1
Jeremy Bentham's Design of the Panopticon

The architecture implements the purpose punishment is supposed to serve (Bentham 1830: Book V Chapter III):

> "Hence the prevention of offenses divides itself into two branches: Particular prevention, which applies to the delinquent himself; and general prevention, which is applicable to all the members of the community without exception.
>
> General prevention is effected by the denunciation of punishment, and by its application, which, according to the common expression, serves for an example. The punishment suffered by the offender presents to every one an example of what he himself will have to suffer if he is guilty of the same offense.
>
> General prevention ought to be the chief end of punishment, as it is its real justification. [...] The punishment inflicted on the individual becomes a source of security to all. [...] That punishment, which, considered in itself, appeared base and repugnant to all generous sentiments, is elevated to the first rank of benefits, when it is regarded not as an act of wrath or of vengeance against a guilty or unfortunate individual who has given way to mischievous inclinations, but as an indispensable sacrifice to the common safety" (Bentham 1830: Book I Chapter III).

1 University College London, Bentham Project and http://en.wikipedia.org/wiki/Panopticon.

This paper investigates Bentham's intuition in a lab experiment. A lab experiment has the typical advantage of full control over the institutional setting. It becomes possible to manipulate how much information on punishment experienced by others is available. Additionally, one can unambiguously see whether and to which degree subjects are well-behaved, and whether or not they change their behavior after having observed or experienced punishment. The experiment is a standard public goods game. It is well known what is to expect, short of the manipulation of this experiment (Ostrom, Walker et al. 1992, Fehr and Gächter 2000, Fischbacher and Gächter 2010). Four players interact over 10 periods. In each period they can contribute to the public good. The individual return of each player from one unit of contribution, i.e. the marginal capita rate, is such that from an individual perspective it is unprofitable to contribute. However, since all players benefit, full contributions from all group maximize total profit. In a second stage of each round, a fifth player observes the individual contributions and, based on this information, can punish each of the four players who can make contributions. Punishing is costly for the fifth player, i.e., she must invest her own money if she decides to diminish the payoff of one or more of the others. All five players have the same return from the public good, irrespective of whether or not and how much they contribute.

To investigate Jeremy Bentham's idea, feedback is manipulated. In the *low* treatment, contributors only learn aggregate contributions. In the *medium* treatment, they also learn aggregate punishment. In the *high* treatment, they know individual contributions and individual punishment. As a further test, another group of four contributors is invited for another 10 periods. The group supervisor stays in office. Before the second group starts playing, graphs inform them about their predecessor's performance. The information about the contributions and, if applicable, about punishment in the first 10 periods is the same as was given to contributors during the first phase. Also the degree of feedback is kept constant across phases. For instance if feedback was *low* in phase 1, the group of successors is not informed about punishment either, neither with respect to their predecessors, nor with respect to other players of the current group.

Compared with the tone that Jeremy Bentham set, the result is non-expected: in the experiment, punishment information is at best immaterial. If bystanders learn both behavior and correctional responses at the individual level, they contribute much less to the public project. However, punishment does not miss its intended effect, neither on those punished themselves, nor on bystanders. Yet the main effect is indirect. Bystanders and newly arrived group members are chiefly influenced by the observed previous contribution levels in their groups that, in turn, have been affected by punishment.

The remainder of the paper is organized as follows: section II relates the paper to the literature and defines the contribution. Section III translates Jeremy Bentham's conjecture into a theoretical claim. Section IV explains the design of the experiment. Section V reports results. Section VI discusses the inevitable limitations inherent in any lab experiment. Section VII concludes.

II. Related Literature

Behavior in public good experiments exhibits a robust and well-known pattern. In the absence of punishment, players contribute significantly in the beginning, but contributions decay quickly (for summaries see Ledyard 1995, Zelmer 2003, Chaudhuri 2011). Contributions stabilize if subjects are given a chance to punish each other after having observed individual contributions (Fehr and Gächter 2000, Fehr and Gächter 2002, Herrmann, Thöni et al. 2008).

Closest to this paper are experiments that manipulate feedback in a public good with punishment. Khadjavi, Lange et al. (2017) implement a game with asymmetric action spaces. While three group members can only contribute (or keep their endowments), one group member can also take from the pool. If individual choices are known, punishment proves more effective. Xiao and Houser (2011) implement automatic punishment. In 50% of all cases, a group is monitored and the lowest contributor is fined. Contributions are higher if the recipient of punishment is made public. This does at least not contradict Jeremy Bentham. Faillo, Grieco et al. (2013) modify the punishment technology. Group members can only punish others if they have contributed less than they themselves. With this manipulation in place, contributions are higher if participants have full feedback, rather than only in the aggregate and of those participants who have contributed less. Patel, Cartwright et al. (2010) have a treatment in which only the identity of participants who have made positive contributions is disclosed before group members decide about punishment. In this treatment, punishment does not stabilize contributions. These results suggest that Jeremy Bentham got it right. Yet none of these experiments directly targets the "general prevention" effect that Jeremy Bentham postulates. This is the contribution of the present paper.

In Ambrus and Greiner (2012) the feedback manipulation is of a different nature: either group members are perfectly informed about the contributions others have made to the public good, or with probability 10% they wrongly learn that a participant has not contributed anything, although she has made a positive contribution. This uncertainty slightly reduces contributions if punishment is particularly strong. Uncertainty has a strong detrimental effect if it is more pronounced (Bornstein and Weisel 2010, Grechenig, Nicklisch et al. 2010, Fischer, Grechenig et al. 2013), and if the admissibility of punishment is subject to group vote (Ambrus and Greiner 2015).

If participants receive feedback about other group members' earnings, contributions are lower than if feedback is about their contributions (Nikiforakis 2010, Bigoni and Suetens 2012).

If there is no punishment, giving participants not only feedback about the aggregate, but also about choices of individual group members in some experiments reduces contributions (Carpenter 2004, Bigoni and Suetens 2012), in other experiments increases contributions (Sell and Wilson 1991, Cox and Stoddard 2015, Kreitmair 2015), and in yet other experiments does not have a significant effect on contributions (Weimann 1994, van der Heijden and Moxnes 1999, Croson 2001), (also see Zylbersztejn 2015).

Selectively informing participants about high contributions only increases contributions if the selection rule is not made transparent (Irlenbusch, Rilke et al. 2018).

In the public good literature, punishment is typically decentralized. Yet increasingly experimental studies use centralized punishment. They for instance investigate whether unaffected outsiders are willing to spend money for disciplining others (Fehr and Fischbacher 2004), how leaders can motivate members of their team by the threat of punishment (Güth, Levati et al. 2007, Gürerk, Irlenbusch et al. 2009), whether centralized punishment develops endogenously when players can voluntarily join a sanctioning institution (Kosfeld, Okada et al. 2009), how centralized punishment affects behavior in a threshold public goods game (Guillén, Schwieren et al. 2006), and how mild central sanctions interact with social norms (Tyran and Feld 2006, Galbiati and Vertova 2008b, Galbiati and Vertova 2008a, Engel 2014).

Very few experimental studies investigate the impact of information about others' behavior (also called "social history") on own behavior. Berg, Dickhaut et al. (1995) find that providing a social history increases cooperation in their trust game setting. Fehr and Rockenbach (2003), however, do not find a change in subjects' behavior in a gift exchange game with punishment. Informing responders about the average offers before they decide whether to accept or reject their specific offer seems to significantly increase offers and offer-specific rejection probabilities (Bohnet and Zeckhauser 2004). In a binary dictator game Krupka and Weber (2009) find that showing subjects what others actually do produces more pro-social behavior. Interestingly, this is even the case when observed subjects are mostly selfish. They also find support for an informational effect: observing more people behaving pro-socially generally produces more pro-social behavior. There does not seem to be a study that investigates the influence of information about others' behavior in a public good setting with punishment.

There is a growing body of experiments in criminology (for summaries see Farrington 2003, Nagin and Pogarsky 2003, Petrosino, Turpin-Petrosino et al. 2003, Farrington and Welsh 2005, Farrington and Welsh 2006, Petrosino, Kiff et al. 2006, Engel 2016b). Many are quasi experiments in the field (Farrington and Welsh 2006). Apparently though no experiment has tried to assess the effect of punishment on true outsiders to the criminal system (cf. the comprehensive survey by Farrington and Welsh (2006) and the survey by Engel (2016a).

III. Hypotheses

The purpose of this experiment is to test the conjecture on which Jeremy Bentham's panopticon proposal builds. It lends itself to formalization. The theoretical framework for this experiment is derived from the observation that sanctions (in the criminal system no less than in the experiment) are meted out by human agents. These agents are fallible. Their reaction function may be noisy, leading to a certain degree of inconsistency and hence unpredictability. This yields

$$E(s_{it}) = f(c_{it}, c^*, \varepsilon) \quad (1)$$

where s is the amount subtracted from participant i's gross profit in period t, $E(s_{it})$ is the active participant's expectation about the authority's punishment policy, c is her actual contribution to the public good, while c^* is the contribution norm the authority wants to impose. Noise ε may result from any of four sources: (a) the participant does not know which precise norm c^* the authority wants to impose; (b) she does not know the authority's reaction function: how does the severity of punishment relate to the intensity of the infraction, i.e. the degree by which the actual contribution is below the desired contribution? (c) how likely is the authority to detect the rule violation, and to react to this information? (d) how consistent is the authority in her punishment choices?

Due to this uncertainty, the would-be criminal maximizes

$$E(u_{it}) = b_{it} - g(E(s_{it}), E(\sigma)) \quad (2)$$

where $E(u_{it})$ is expected utility. The utilitarian criminal trades the benefit b from committing the crime against her sensitivity $g(.)$ towards the risk of being punished. For Bentham's argument it is critical that the individual is not only sensitive to the first moment of the distribution of potential sanctions (i.e. the expected value of the sanction $E(s_{it})$), but also to the second moment of this distribution (i.e. the expected variance $E(\sigma)$). One may also interpret σ as the perceived precision of the authority's reaction function. Bentham's thinking implies

$$E(\sigma_{high}) < E(\sigma_{medium}) < E(\sigma_{low})$$

where $high, medium, low$ stand for the treatments of the experiment. Jeremy Bentham's claim further implies: the more the information about the punishment function of the authority is precise, the more the individual is deterred: $\frac{\partial g}{\partial \sigma} < 0$.

Information about the way in which the authority has reacted to foreign contribution choices is as informative about her sanction policy as are the experiences the participant has made herself. This holds for all four sources of uncertainty (a) - (d). In the *low* treatment the participant has no direct information about punishment meted out to other participants. The only signal is (the development of) her own profit, as profit depends on the contributions made by other participants. If punishment induces them to increase their contributions, the participant sees the effect in her own period profit. By contrast in the *medium* treatment, she has explicit information about the choices of others and the punishment they have received. Yet as she only learns aggregates, this information is not very precise. In the *high* treatment, the information about the past is perfect. The only potentially remaining source of uncertainty is inconsistency in the punishment policy of the respective authority (d). Yet as group composition stays constant, in the *high* treatment the participant even has information about actual variance in the authority's punishment choices.

The theoretical framework implies:

H_1: Contributions are highest in the *high* treatment, lower in the *medium* treatment, and *lowest* in the low treatment.

H_2: a) Participants increase their contributions the more the more often and the more severely they have been punished themselves.
b) In the *medium* and *high* treatments, participants also increase their contributions the more often and the more intensely other group members have been punished.

H_3: In the *high* treatment, participants react more intensely to punishment received by other participants than in the *medium* treatment

IV. Experimental Design

Participants play announced 10 rounds of a standard public goods game in anonymous groups of four (called "players of type A" in the instructions) that stay together during the entire experiment. Per period, each participant receives an endowment of 20 talers. Players of type A can decide how many talers to invest in a project. Each taler contributed to the project creates a marginal per capita return of 0.4. Hence gross profit is given by (3)

$$\pi_{it} = 20 - c_{it} + .4\sum_{k=1}^{4} c_{kt} \quad (3)$$

where k is generic for each of the 4 members of the group.

Before the start of the game, per group one additional subject is randomly assigned as the group supervisor ("player of type B" in the instructions). In each round supervisors are informed about the individual contributions of each type A-player. Supervisors have the same endowment and the same marginal per capita return from the project. However, they cannot contribute. Rather they can spend their endowment on individually punishing the type A-players. For type B players, contributions of type A players thus are the equivalent of a levy from which a public official is financed. The punishment technology is linear: one taler invested for punishment destroys three talers of the punished player. Hence net profit is given by (4)

$$\pi_{it} = 20 - c_{it} + .4\sum_{k=1}^{4} c_{kt} - 3 * p_{it} \quad (4)$$

where p stands for each punishment point allotted to this player in this period.

The experiment consists of two phases. The first group of four type A players in phase 1 is followed by a second phase with a fresh group of four subjects. Also the second phase consists of announced 10 rounds. Only the supervisor stays the same in both phases. Before starting to play themselves, the second group receives graphs informing them about the performance and/or received punishment in their respective group of predecessors.

The three treatments differ in feedback. An overview of the differences in feedback is provided in Table 1.

Treatment	feedback for supervisor	feedback for active players in phase 1	additional feedback for active players in phase 2 about phase 1
low	• individual contributions, players not identified across periods	• average contributions • own received punishment	• average contributions
medium (in addition to information provided in *low*)	• individual contributions, players not identified across periods	• average received punishment	• average received punishment
high (in addition to information provided in *medium*)	• individual contributions, players identified across periods	• individual contributions • individual earnings • individual received punishment	• individual contributions • individual earnings • individual received punishment

Table 1

Feedback Provided to Type A-Players in Different Treatments

Participants receive a show up fee of 2.50 €. Theoretically, subjects can make real losses. Since the lab has built a reputation that subjects do not put their own money at risk, they receive an extra 50 talers at the beginning of the experiment, explicitly motivated to cater for potential losses. Earnings are individually and anonymously paid out to all participants at an exchange rate of

0.04 € per taler. On average, total earnings of contributors were 22.23 € (sd 2.64, range [12.07, 33.3]). Total earnings of supervisors, who played 20 periods each, were on average 45.60 € (standard deviation 5.51, range [34.36, 56.71]).

324 students (149 female) from a variety of majors participated in the experiment conducted at the Econ Lab of Cologne University. The experiment was implemented in zTree (Fischbacher 2007). Participants were invited using Orsee (Greiner 2004) and were randomly assigned to treatments.[2]

V. Results

1. Anticipation

In the first period of the first phase, active players have no own or vicarious experiences with the respective punishment institution. But they are fully informed about institutional design. Is this information sufficient to induce different behavior in different treatments? Is a potential effect of the institution anticipated? Although descriptive figures point into the direction of the treatment effects[3], in the first round there are no significant treatment effects, neither non-parametrically[4] nor parametrically.[5]

2. Phase 1

a. *The Effect of Transparency on Contributions*

Contributions and deductions through punishment are as in Figure 2 and in Table 2. The main result is patent: full transparency hurts, while partial transparency is immaterial. In treatment *high*, absolute contributions are lower than in the remaining two treatments. The difference in absolute contributions results from less favorable contribution dynamics. While contributions rise quickly in treatments *low* and *medium*, they remain almost stable in treatment *high*. Visual inspection suggests that informing participants about average punishment, i.e. the difference between treatments *low* and *medium*, is close to irrelevant.

Phase	first		second	
Treatment	contribution	punishment	contribution	punishment
Low	17.45	1.7	18.07	1.65
Medium	16.99	2.68	18.18	1.72
High	14.02	2.84	15	1.85

Table 2

Means of Contributions and Received Punishment

2 The translation of the instructions for one of our treatments are provided in the appendix. Original instructions were in German. All instructions can be obtained upon request.

3 Mean contributions are low 14.521, medium 14.563, high 13.208.

4 Mann Whitney, low vs. medium, $N = 96$, $p = .932$; low vs. high, $p = .2955$; medium vs. high, $p = .2514$. All tests are two-sided. Note that in the first period individual contribution decisions are still fully independent of each other.

5 In the first period, 61 out of 144 (active) participants contribute their entire endowment of 20 taler. 4 keep the whole endowment for themselves. This data structure makes a Tobit model appropriate. In this model, treatment high is the reference category. Regressors for treatments low ($p = .229$) and medium ($p = .184$) are independently and jointly insignificant (Wald test, $F(2, 142) = 1.09$, $p = .3376$).

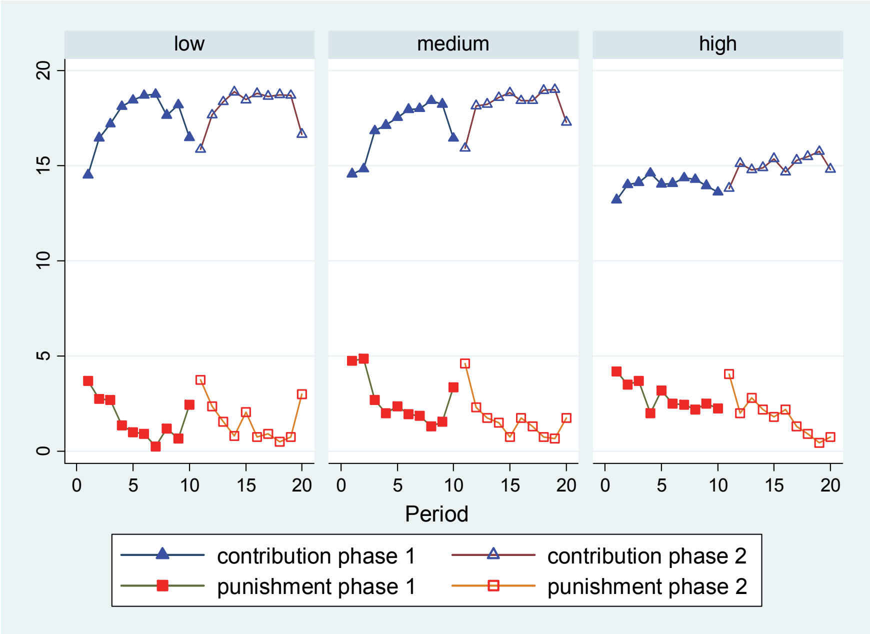


Figure 2
Descriptives

Non-parametrically, the difference between *low* and *high* is weakly significant (Mann-Whitney over mean contributions per group, N = 24, p = .0647), while the remaining comparisons are insignificant. Parametric estimation captures the nested character of the data (choices in individuals in groups), as well as upper and lower censoring, and controls for the pronounced end game effect. Using this strategy, one establishes significantly higher contributions in the *low* and *medium* treatments, compared to treatment *high*, which serves as the reference category (Table 3). Full feedback (*high*) thus reduces contributions. This squarely refutes $\mathbf{H_1}$.

low	7.601* (3.176)
medium	6.794* (3.184)
final period	-1.256^{+} (.737)
cons	16.831*** (2.234)
N	1440
left censored	52
right censored	784

Table 3
Parametric Test of Treatment Effects
depvar: contribution
final period: a dummy that is 1 in period 10
mixed effects Tobit
standard errors for choices nested in individuals nested in groups in parenthesis
*** p < .001, ** p < .01, * p < .05, + p < .1

Making average punishment explicit (*medium*) does not significantly improve contributions, as shown by a Wald test of the null hypothesis that coefficients for *low* and *medium* are the same ($p = .8008$). This yields

> *Result 1:* In a linear public good, contributions are lower if participants have information about the punishment of other individual group members.

b. Sensitivity Towards the Experience of Being Punished

As the regressions in Table 4 show, participants increase their contributions when they have been punished in the previous period (model 1), the more so the more severely they have been punished (model 2). We thus have full support for $\mathbf{H_{2a}}$ and formulate

> *Result 2*: In a linear public good, participants increase their contributions if they have been punished in the previous period.

Yet in square contradiction to hypothesis $\mathbf{H_{2b}}$, the more severely the remaining group members have been punished, *the less* participants increase their own contributions. Seeing others punished does not help but hurt, even if this is only aggregate information. This result contradicts a first intuition on which Jeremy Bentham's thinking is based. The contradiction is plain in model 4. If one interacts information about own and foreign punishment with treatment *high*, the interaction is insignificant for foreign punishment: there is no support for the critical piece of Jeremy Bentham's claim, which is expressed in $\mathbf{H_{3}}$. Actually, the interaction between this treatment and own punishment is even significantly negative. If punishment is fully transparent, participants are even less sensitive to the experience of having been punished themselves. This yields

> *Result 3*: The more other group members in a linear public good have on average been punished in the previous period, the more others reduce their contributions. This negative effect is most pronounced if they learn how much others have been punished individually.

	model 1	model 2	model 3	model 4
high				.010 (.319)
punished_{t-1}	3.731*** (.272)			
punishment_{t-1}		.752*** (.054)	.578*** (.054)	.693*** (.082)
*high**punishment_{t-1}				-.220* (.110)
mean punishment of others_{t-1}			-.088** (.029)	-.085^{+} (.051)
*high** mean punishment of others_{t-1}				-.002 (.062)
cons	-.779*** (.160)	-.439** (.142)	-.164 (.155)	-.171 (.238)
N	1296	1296	864	864

Table 4
Reactions to Punishment
depvar: contributiont – contributiont-1
models 1-2: data from all treatments
models 3-4: data from treatments medium and high
(as this information was not available in treatment low).
linear mixed effects
standard errors for choices nested in individuals nested in groups in parenthesis
*** p < .001, ** p < .01, * p < .05, + p < .1

c. *The Missing Link*

What did Jeremy Bentham get wrong? Table 5 provides the missing link. It reports a structural model that simultaneously estimates the determinants of contributions, and the determinants of these determinants. Participants do respond to information about the punishment of others, in the direction Jeremy Bentham expected: if others have been punished more, they contribute more themselves. But they also respond to information about the choices of others. If all others have contributed one Taler more in the previous period, they contribute half a Taler more themselves. Others may at most contribute 60 Taler. They on average contribute 46.688 Taler. They may at most have received 20 punishment points. On average they have received 2.753 Taler. Taken these distributions of the explanatory variables into account, the regression shows that participants react much more to foreign contributions than to foreign punishment. They are less interested in wrongdoing being punished, and more interested in socially acceptable behavior of others being achieved.

The significant negative interaction term shows a further effect. If high contributions can only be achieved with high punishment, participants contribute less themselves. In the best of all worlds, others behave well, with not much need of enforcement.

The most important piece of the puzzle is, however, contained in the remaining two components of the structural model. In the *high* treatments, other group members have on average been punished more severely in the previous period (third component), but they have contributed much less (second component). Full transparency starts a vicious cycle. Of necessity, bystanders not only learn how determined the authority is to keep misbehavior in check; they also learn how poorly some others behave. The fact that others try to exploit them is no longer concealed in the average. They observe each individual instance of exploitation.

contribution	
total contribution of others$_{t-1}$	.497*** (.027)
total punishment of others$_{t-1}$	.761** (.247)
total contribution of others$_{t-1}$ * total punishment of others$_{t-1}$	-.025*** (.006)
cons	-3.364* (1.405)
total contribution of others$_{t-1}$	
high	-6.947*** (1.110)
cons	50.161*** (.920)
total punishment of others$_{t-1}$	
high	2.323*** (.579)
cons	1.592* (.631)
N	864

Table 5
Determinants of Contributions
depvar: contribution
mixed effects structural model, first component Tobit, other two components linear
standard errors for choices nested in individuals nested in groups in parenthesis
*** p < .001, ** p < .01, * p < .05, + p < .1

This yields the final

> *Result 4*: In a linear public good, participants react more intensely to information about the past contributions of others than about the past punishment of others.

The conjecture of Jeremy Bentham can be theoretically captured by sensitivity of choices to the expected variability of punishment (section III). As soon as participants have experience, they can use them to update their homegrown expectations. If Jeremy Bentham gets it right, in treatment *high*, where participants have this information, contributions should be lower the more the reaction of the authority to the observed level of contributions has been erratic. In the regression of Table 6, variability is captured by the standard error of the coefficient of contribution in a local regression of punishment on contributions, separately for each group and for the complete past. In Table 6, the effect of this standard error actually even has the "wrong" sign: the more punishment has been variable, the more (not the less) participants contribute. This shows that the effect on which Jeremy Bentham's thinking is based is clearly not present in the data.

severity of punishment until $_{t-1}$	.336 (7.962)
variability of punishment until $_{t-1}$	80.454** (28.886)
cons	14.494*** (3.360)
N	384

Table 6

Sensitivity of Choices to Severity and Variability of Punishment
depvar: contribution
severity of punishment is the coefficient of a local regression
of punishment on contribution, for periods 2-(t-1), separately for each group
variability of punishment is the coefficient of a local regression
of punishment on contribution, for periods 2-(t-1), separately for each group
data from phase 1, and treatment high
mixed effects Tobit
standard errors for choices nested in individuals nested in groups in parenthesis
*** p < .001, ** p < .01, * p < .05

3. Foreign Experiences

The data from the second phase of the experiment fit the picture. As Figure 2 shows, descriptively the data look very similar to the first phase. This impression is supported by the regression in Table 7. At the beginning of the second phase the new group of active participants receives graphical information about choices and (in treatments *medium* and *high*) punishment in the first phase in their group. In the statistical model of Table 7, this information is captured by a local regression that, separately for each group, regresses both contributions and punishment on period. A positive coefficient means that contributions or punishment have been increasing over time. A negative coefficient means that they have been decreasing. As Table 7 shows, participants do react to this information. If contributions in the predecessor group have been increasing over time, they contribute more. If punishment in the predecessor group has been increasing over time, they contribute less. Conditional on these foreign experiences, they contribute more the more their own group members have contributed in the previous period. Unlike the first phase, participants in the second phase also contribute more if other members of their group have been punished more severely in the previous period. This difference likely results from the fact that the negative effect of high punishment is already captured by the experiences from the previous phase.

development of contributions in phase 1	4.197* (2.016)
development of punishment in phase 1	-14.848* (5.261)
total contribution of others$_{t-1}$	.240*** (.051)
total punishment of others$_{t-1}$	.293* (.138)
cons	4.172 (2.973)
N	864

Table 7

Own and Foreign Experiences

depvar: contribution

development of contributions in phase 1 is the coefficient of a local regression of contributions on period, for periods 1-10, separately for each group

development of punishment in phase 1 is the coefficient of a local regression of punishment on period, for periods 1-10, separately for each group

data from phase 2, and treatments medium and high

mixed effects Tobit

standard errors for choices nested in individuals nested in groups in parenthesis

*** p < .001, ** p < .01, * p < .05

VI. Discussion

Lab experiments are not meant, and they are not able, to fully capture the richness of the real life phenomenon that motivates the endeavor. This is not a bug, but a feature. Precisely because the experiment abstracts from all other elements, it is able to causally identify the effect of interest. In this experiment, it is the effect of transparency about the punishment of others on the choices of bystanders. This section discusses in which ways this result informs the policy debate in the field.

In the reality of criminal policy, stakes are much higher, both for society and for the "offender". But the experiment keeps the basic dilemma structure that also underlies most criminal offenses: the criminal is best off if she ignores the harm she inflicts on other members of society.

There is no criminal code. Norms are implicit. They result from the sanctioning policy of the supervisor. Yet in criminal policy, a related effect is not uncommon: the criminal authorities use the degrees of freedom they dispose of to flexibly react to crime, despite the fact that, at face value, criminal offenses are precisely defined in the respective penal code.

In the experiment, sanctions are not accompanied by words that express social disapproval or moral indignation. In the experiment, all value judgement is through correctional action. The effect of the negative incentive is less visibly backed up by appealing to the offender's identity as a member of this one society.

Supervisors cannot personally identify group members so that repeated game effects are not an issue. Effects of retribution can be excluded (cf. Wood 2002). Incapacitation is impossible (cf.

Kessler and Levitt 1999). Intervention cannot shift criminal activity to another location (cf. Hakim, Spiegel et al. 1984). Victims cannot respond by moving to a different town or quarter (cf. Anderson 1990). Players are perfectly symmetric, so that bystanders have no reason to expect that they will be treated any differently if they behave the same way (cf. Robinson and Darley 2003: 973). Since each round only lasts minutes, arguably discounting should be negligible (cf. Levitt 1998: 353).

Given the completely neutral framing and decontextualisation and the fact that there cannot be competing tasks, impulsivity (cf. Shepherd 2004) should not play a role, nor crime as symbol (cf. Matsueda, Kreager et al. 2006: 103). The risk of sanction misperception (Nagin 1998: 19) is minimized. Subjects are perfectly informed about punishment inflicted on themselves and, depending on our treatments, on others. Arguably, habituation does not matter, given that players only play 10 rounds, and that the entire experiment lasts little longer than an hour (cf. Hawkins 1969: 560). Again due to anonymity, the fact that would-be offenders are members of a peer group with criminal propensity cannot explain behavior (cf. Kahan 1997a: 2486). Punishment cannot serve as a "badge of honor" (cf. Wilson and Herrnstein 1985: 304). Moral credibility (cf. Kahan 1997a: 2481) and the mirror concept of moral condemnation (cf. Kahan 1997b: 383) cannot matter either. Since anonymity is guaranteed, formal sanctions cannot be supplemented or complemented by informal sanctions in treatments *low* and *medium* (cf. Cameron 1988: 302). In treatment *high*, free-riders are labeled (cf. Lemert 1951, Becker 1963). But other group members do not have an opportunity for a targeted reaction.

VII. Conclusions

Jeremy Bentham had the conjecture that punishment must be made transparent if it is to guide those who have not been punished themselves. This conjecture can be backed by theory. The result follows if individuals tempted to misbehave are sensitive to the predictability of sanctions. This hypothesis is, however, not supported by the experimental evidence. If contributions and punishment are transparent, the willingness to contribute to the common cause decays. Even more disturbingly: actual sanctions become less effective the better would-be low contributors can observe how the criminal system reacts to offenses. The results suggest that instead of building panopticon prisons in the town centre, government should conceal prisons from public scrutiny. Punishment serves society best if it remains a tool one does not see in action when applied to others. What really matters is information about normabiding behavior of other members of society. To use a metaphor that features prominently in criminal policy: government should spend money on repairing broken windows, not on showcasing correctional action.

Nonetheless, this is a paper in the spirit of Jeremy Bentham. While he might have got it wrong as a policy maker, he got it totally right as an analyst. The main task of criminal policy is and ought to be that would-be criminals are induced to exhibit socially desirable behavior. Only the route to the end is a different one. The main tool ought to be impression management, not deterrence.

References

AMBRUS, ATTILA AND BEN GREINER (2012). "Imperfect Public Monitoring with Costly Punishment. An Experimental Study." American Economic Review **102**(7): 3317-3332.

AMBRUS, ATTILA AND BEN GREINER (2015). Democratic Punishment in Public Good Games with Perfect and Imperfect Observability. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2567326.

ANDERSON, ELIJAH (1990). Streetwise. Race, Class, and Change in an Urban Community. Chicago, University of Chicago Press.

BECKER, HOWARD SAUL (1963). Outsiders. Studies in the Sociology of Deviance. London,, Free Press of Glencoe.

BENTHAM, JEREMY (1830). The Rationale of Punishment. London,, R. Heward.

BERG, JOYCE, JOHN DICKHAUT AND KEVIN MCCABE (1995). "Trust, Reciprocity, and Social History." Games and Economic Behavior **10**: 122-142.

BIGONI, MARIA AND SIGRID SUETENS (2012). "Feedback and Dynamics in Public Good Experiments." Journal of Economic Behavior & Organization **82**(1): 86-95.

BOHNET, IRIS AND RICHARD ZECKHAUSER (2004). "Social Comparisons in Ultimatum Bargaining." Scandinavian Journal of Economics **106**: 495-510.

BORNSTEIN, GARY AND ORI WEISEL (2010). "Punishment, Cooperation, and Cheater Detection in "Noisy" Social Exchange." Games **1**(1): 18-33.

CAMERON, SAMUEL (1988). "The Economics of Crime Deterrence. A Survey of Theory and Evidence." Kyklos **41**: 301-323.

CARPENTER, JEFFREY P (2004). "When in Rome. Conformity and the Provision of Public Goods." Journal of Socio-Economics **33**(4): 395-408.

CHAUDHURI, ANANISH (2011). "Sustaining Cooperation in Laboratory Public Goods Experiments. A Selective Survey of the Literature." Experimental Economics **14**(1): 47-83.

COX, CALEB A AND BROCK STODDARD (2015). "Framing and Feedback in Social Dilemmas with Partners and Strangers." Games **6**(4): 394-412.

CROSON, RACHEL TA (2001). Feedback in Voluntary Contribution Mechanisms. An Experiment in Team Production. Research in Experimental Economicss, Emerald Group Publishing Limited**:** 85-97.

ENGEL, CHRISTOPH (2014). "Social Preferences Can Make Imperfect Sanctions Work. Evidence from a Public Good Experiment." Journal of Economic Behavior & Organization **108**: 343-353.

ENGEL, CHRISTOPH (2016a). Experimental Criminal Law. A Survey of Contributions from Law, Economics and Criminology. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2769771.

ENGEL, CHRISTOPH (2016b). Experimental Criminal Law. A Survey of Contributions from Law, Economics and Criminology. EE Handbook on Empirical Methods in Legal Research. W. v. Boom, P. T. Desmet and P. Mascini**:** ***.

FAILLO, MARCO, DANIELA GRIECO AND LUCA ZARRI (2013). "Legitimate Punishment, Feedback, and the Enforcement of Cooperation." Games and Economic Behavior **77**(1): 271-283.

FARRINGTON, DAVID P. (2003). "A Short History of Randomized Experiments in Criminology. A Meager Feast." Evaluation Review **27**: 218-227.

FARRINGTON, DAVID P. AND BRANDON C. WELSH (2005). "Randomized Experiments in Criminology. What Have We Learned in the Last Two Decades?" Journal of Experimental Criminology **1**: 9-38.

FARRINGTON, DAVID P. AND BRANDON C. WELSH (2006). "A Half Century of Randomized Experiments on Crime and Justice." Crime and Justice **34**: 55-132.

FEHR, ERNST AND URS FISCHBACHER (2004). "Third-Party Punishment and Social Norms." Evolution and Human Behavior **25**: 63-87.

FEHR, ERNST AND SIMON GÄCHTER (2000). "Cooperation and Punishment in Public Goods Experiments." American Economic Review **90**: 980-994.

FEHR, ERNST AND SIMON GÄCHTER (2002). "Altruistic Punishment in Humans." Nature **415**: 137-140.

FEHR, ERNST AND BETTINA ROCKENBACH (2003). "Detrimental Effects of Sanctions on Human Altruism." Nature **422**: 137-140.

FISCHBACHER, URS (2007). "z-Tree. Zurich Toolbox for Ready-made Economic Experiments." Experimental Economics **10**: 171-178.

FISCHBACHER, URS AND SIMON GÄCHTER (2010). "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Good Experiments." American Economic Review **100**: 541-556.

FISCHER, SVEN, KRISTOFFEL R GRECHENIG AND NICOLAS MEIER (2013). Cooperation under Punishment. Imperfect Information Destroys it and Centralizing Punishment Does Not Help. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2243478.

GALBIATI, ROBERTO AND PIETRO VERTOVA (2008a). How Laws Affect Behaviour. http://ssrn.com/abstract=1295948.

GALBIATI, ROBERTO AND PIETRO VERTOVA (2008b). "Law and Behaviours in Social Dilemmas. Testing the Effect of Obligations on Commitment." Games and Economic Behavior **64**: 146-170.

GRECHENIG, KRISTOFFEL, ANDREAS NICKLISCH AND CHRISTIAN THÖNI (2010). "Punishment Despite Reasonable Doubt. A Public Goods Experiment with Sanctions Under Uncertainty." Journal of Empirical Legal Studies **7**(4): 847-867.

GREINER, BEN (2004). An Online Recruiting System for Economic Experiments. Forschung und wissenschaftliches Rechnen 2003. K. Kremer and V. Macho. Göttingen, GWDG**:** 79-93.

GUILLÉN, PABLO, CHRISTIANE SCHWIEREN AND GIANANDREA STAFFIERO (2006). "Why Feed the Leviathan?" Public Choice **130**: 115-128.

GÜRERK, ÖZGÜR, BERND IRLENBUSCH AND BETTINA ROCKENBACH (2009). "Motivating Teammates. The Leader's Choice of Positive and Negative Incentives." Journal of Economic Psychology **30**: 591-607.

GÜTH, WERNER, VITTORIA M. LEVATI, MATTHIAS SUTTER AND ELINE VAN DER HEIJDEN (2007). "Leading by Example With and Without Exclusion Power in Voluntary Contribution Experiments." Journal of Public Economics **91**: 1023-1042.

HAKIM, SIMON, URIEL SPIEGEL AND J. WEINBLATT (1984). "Substitution, Size Effects, and the Composition of Property Crime." Social Science Quarterly **65**: 719-734.

HAWKINS, GORDON (1969). "Punishment and Deterrence. The Educative, Moralizing, and Habituative Effects." Wisconsin Law Review: 550-565.

HERRMANN, BENEDIKT, CHRISTIAN THÖNI AND SIMON GÄCHTER (2008). "Antisocial Punishment Across Societies." Science **319**: 1362-1367.

IRLENBUSCH, BERND, RAINER MICHAEL RILKE AND GARI WALKOWITZ (2018). Designing Feedback in Voluntary Contribution Games. The Role of Transparency. https://nbn-resolving.org/urn:nbn:de:hbz:992-opus4-6396.

KAHAN, DAN (1997a). "Between Economics and Sociology. The New Path of Deterrence." Michigan Law Review **95**: 2477-2497.

KAHAN, DAN (1997b). "Social Influence, Social Meaning, and Deterrence." Virginia Law Review **83**: 349-395.

KESSLER, DANIEL AND STEVEN D. LEVITT (1999). "Using Sentence Enhancements to Distinguish Between Deterrence and Incapacitation." Journal of Law and Economics **42**: 343-363.

KHADJAVI, MENUSCH, ANDREAS LANGE AND ANDREAS NICKLISCH (2017). "How Transparency May Corrupt. Experimental Evidence from Asymmetric Public Goods Games." Journal of Economic Behavior & Organization **142**: 468-481.

KOSFELD, MICHAEL, AKIRA OKADA AND ARNO RIEDL (2009). "Institution Formation in Public Goods Games." American Economic Review **99**: 1335-1355.

KREITMAIR, URSULA W (2015). "Voluntary Disclosure of Contributions. An Experimental Study on Nonmandatory Approaches for Improving Public Good Provision." Ecology and Society **20**(4): 33.

KRUPKA, ERIN AND ROBERTO A. WEBER (2009). "The Focusing and Informational Effects of Norms on Pro-Social Behavior." Journal of Economic Psychology **30**: 307-320.

LEDYARD, JOHN O. (1995). Public Goods. A Survey of Experimental Research. The Handbook of Experimental Economics. J. H. Kagel and A. E. Roth. Princeton, NJ, Princeton University Press**:** 111-194.

LEMERT, EDWIN MCCARTHY (1951). Social Pathology. A Systematic Approach to the Theory of Sociopathic Behavior. New York,, McGraw-Hill.

LEVITT, STEVEN D. (1998). "Why Do Increased Arrest Rates Appear to Reduce Crime. Deterrence, Incapacitation, or Measurement Error?" Economic Inquiry **36**: 353-372.

MATSUEDA, ROSS L., DEREK A. KREAGER AND DAVID HUIZINGA (2006). "Deterring Delinquents. A Rational Choice Model of Theft and Violence." American Sociological Review **71**: 95-122.

NAGIN, DANIEL AND GREG POGARSKY (2003). "An Experimental Investigation of Deterrence. Cheating, Self-Serving Bias, and Impulsivity." Criminology **41**: 167-193.

NAGIN, DANIEL S. (1998). "Criminal Deterrence Research at the Outset of the Twenty-First Century." Crime and Justice **23**: 1-42.

NIKIFORAKIS, NIKOS (2010). "Feedback, Punishment and Cooperation in Public Good Experiments." Games and Economic Behavior **68**(2): 689-702.

OSTROM, ELINOR, JAMES M. WALKER AND ROY GARDNER (1992). "Covenants with and without Sword. Self-Governance is Possible." American Political Science Review **40**: 309-317.

PATEL, AMRISH, EDWARD CARTWRIGHT AND MARK VAN VUGT (2010). Punishment Cannot Sustain Cooperation in a Public Good Game with Free-rider Anonymity. https://gupea.ub.gu.se/handle/2077/22373.

PETROSINO, ANTHONY, PAUL KIFF AND JULIA LAVENBERG (2006). "Randomized Field Experiments Published in the British Journal of Criminology, 1960-2004." Journal of Experimental Criminology **2**: 99-111.

PETROSINO, ANTHONY, CAROLYN TURPIN-PETROSINO AND JOHN BUEHLER (2003). "Scared Straight and Other Juvenile Awareness Programs for Preventing Juvenile Delinquency. A Systematic Review of the Randomized Experimental Evidence." Annals of the American Academy of Political and Social Science **589**(41-62).

ROBINSON, PAUL H. AND JOHN M. DARLEY (2003). "The Role of Deterrence in the Formulation of Criminal Rules. At Its Worst When Doings Its Best." Georgetown Law Journal **91**: 949-1002.

SELL, JANE AND RICK K WILSON (1991). "Levels of Information and Contributions to Public Goods." Social Forces **70**(1): 107-124.

SEMPLE, JANET (1993). Bentham's Prison. A Study of the Panopticon Penitentiary. Oxford, Clarendon.

SHEPHERD, JOANNA M. (2004). "Murders of Passion, Execution Delays, and the Deterrence of Capital Punishment." Journal of Legal Studies **33**: 283-321.

TYRAN, JEAN-ROBERT AND LARS P. FELD (2006). "Achieving Compliance when Legal Sanctions are Non-Deterrent." Scandinavian Journal of Economics **108**: 135-156.

VAN DER HEIJDEN, ELINE CM AND ERLING MOXNES (1999). Information Feedback in Public-bad Games. A Cross-country Experiment. https://ideas.repec.org/p/tiu/tiucen/24bc7f13-1382-47d1-95c7-77d03fb4b345.html.

WEIMANN, JOACHIM (1994). "Individual Behaviour in a Free Riding Experiment." Journal of Public Economics **54**(2): 185-200.

WILSON, JAMES Q. AND RICHARD J. HERRNSTEIN (1985). Crime and Human Nature. New York, Simon and Schuster.

WOOD, DAVID (2002). "Retribution, Crime Reduction and the Justification of Punishment." Oxford Journal of Legal Studies **22**: 301-321.

XIAO, ERTE AND DANIEL HOUSER (2011). "Punish in Public." Journal of Public Economics **95**(7): 1006-1017.

ZELMER, JENNIFER (2003). "Linear Public Goods. A Meta-Analysis." Experimental Economics **6**: 299-310.

ZYLBERSZTEJN, ADAM (2015). Nonverbal Feedback, Strategic Signaling, and Nonmonetary Sanctioning. New Experimental Evidence from a Public Goods Game. Replication in Experimental Economics, Emerald Group Publishing Limited: 153-181.

Appendix

Instructions
(Treatment *High* Second Phase)

(Instructions for phase 1 and for phase 2, treatments *low* and *medium* are available upon request)

General Instructions for Participants

You are about to take part in an economics experiment. If you read the following instructions carefully, you will be able to earn a substantial sum of money, depending on the decisions you make. It is therefore very important that you read these instructions carefully.

The instructions you have received are exclusively for your private information. **There shall be absolutely no communication during the experiment.** If you have questions, please ask us. Disobeying this rule will lead to exclusion from the experiment and any payments.

The experiment consists of several parts. We will begin by explaining the first part. You will receive separate instructions for the other parts.

You will definitely receive € 2.50 for participating in the experiment. During the experiment, the currency in operation is not euro, but taler. Your entire income is hence first calculated in taler. The total number of taler you will have accumulated in the course of the experiment will then be transferred into euro at the following rate:

1 taler = 3 euro cent.

At the end of the experiment, you will receive a **cash** payment, in euro, of whatever number of taler you have earned.

Participants are divided into groups of five. In other words, there are 4 further participants in your group.
All five participants in your group are taking part in this experiment for the first time. There are two roles: four participants, who have confirmed their presence at this experiment for two hours, are assigned Role A. Another person, who has confirmed his presence at this experiment for 4 hours, is assigned Role B.

The experiment is divided into individual periods, of which there are a total of 10. During these 10 periods, the constellation of your group of five remains unchanged. **You are therefore in the same group with the same people for 10 periods. During these 10 periods, the role you have been assigned also remains unchanged.**

At the beginning of the experiment, each participant is given a lump-sum payment of 50 taler. This occurs only once. You may cover possible losses with these 50 taler.

The following pages give you an outline of the exact proceedings of the experiment.

Information on the Exact Proceedings of the Experiment

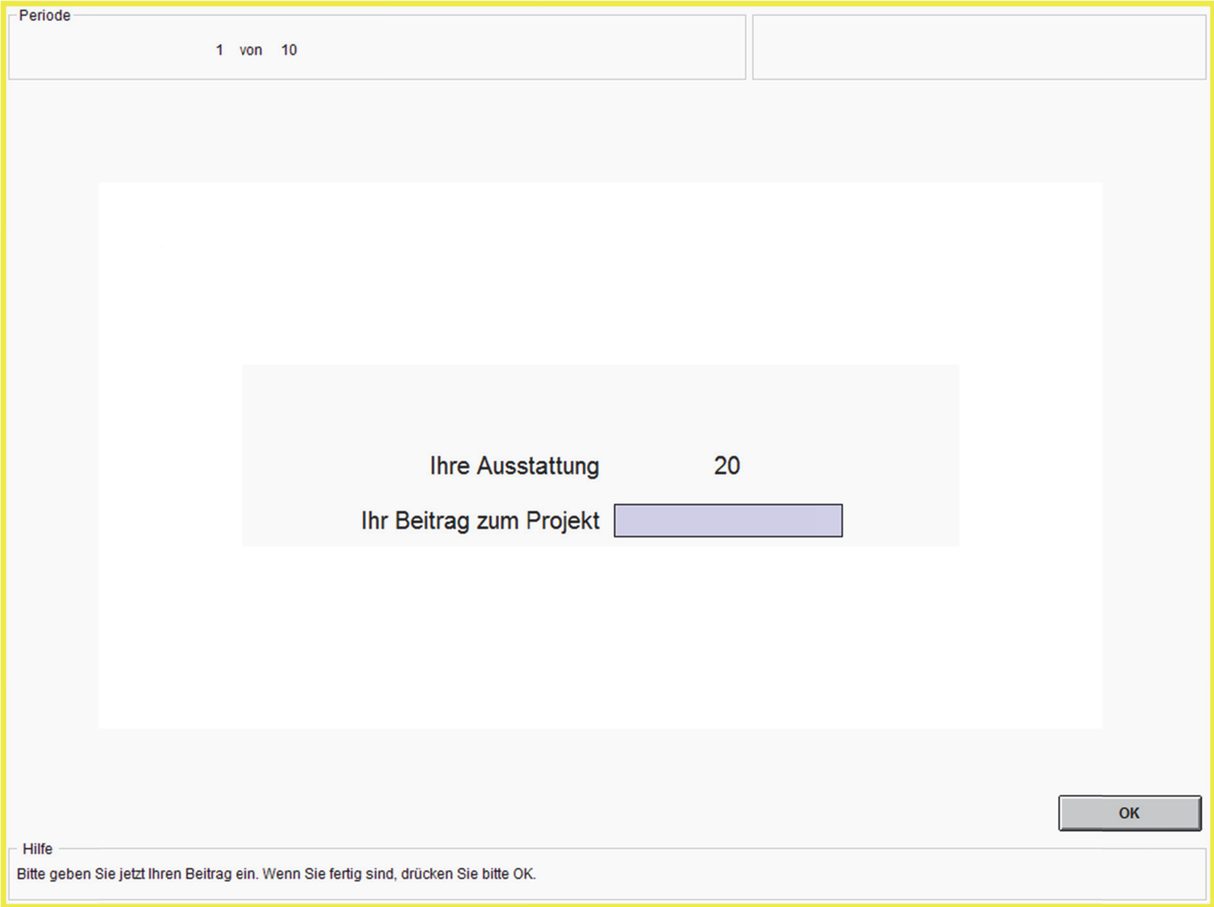
Each of the 10 periods consists of two steps. In Step 1, the participants who have been assigned Role A decide on contributions to a project. In Step 2, the participant who has Role B can reduce the income of the other (Role A) participants. At the beginning of each period, each participant receives **20 points**, referred to henceforth as **endowment**.

Step 1:

In Step 1, **only the four Role A participants** in a group make a decision (should you have been assigned Role B, please read this part of the instructions anyway, in order to find out how a Role A participant can reach a decision). Your task is to reach a decision on how to use your endowment. As a Role A participant, you have to decide how many of the 20 points you wish to pay into a project, and how many you wish to keep for yourself. The consequences of this decision are explained in more detail below.

At the beginning of each period, the following input screen appears:

The input screen:



On the top left corner of your screen, the **period number** is displayed.

As already mentioned, your **endowment in each period is 20 points**. As a Role A participant, you have to make a decision on your project contribution by typing in a sum between 0 and 20 in the appropriate field. You may click on this window by using the mouse. As soon as you have determined the sum you wish to contribute, you have also decided on how many points you keep for yourself: **20 minus your contribution**. Once you have keyed in your amount, press or click **O.K.**, using the mouse or the Enter-key. As soon as you have done this, you can no longer make any changes to your decision.

Your income from the contribution phase consists of two parts:

(1) the points that you have kept for yourself (**"income from endowment kept")**

(2) the **"income from the project"**. The income from the project is calculated as follows:

> Your income from the project =
> 0.4 *times* the total sum of contributions to the project

Your **income from the project, in taler,** for one period is therefore

(20 minus your contribution to the project) + 0.4* (total sum of contributions to the project).

The income of all other group members is calculated according to the same formula, i.e., each group member receives the same income from the project. If, for example, the sum of contributions from all group members is 60 points, then you and all other group members will receive a points income from the project of 0.4*60 = 24 taler. If the group members have contributed a total of 9 points to the project, you and all other group members will receive 0.4*9 = 3.6 taler as your income from the project.

For each point that you keep for yourself, you earn an income of 1 taler. If instead you contribute one point from your endowment to your group project, then the sum of contributions to the project increases by 1 point, and your income from the project increases by 0.4*1 = 0.4 taler. However, this also means that the income of all other group members increases by 0.4 taler, so that the total income of the group increases by 0.4*5 = 2 taler. Through your contributions to the project, the other group members also increase their earnings. On the other hand, you also earn something from the contributions of the other group members to the project. For each point that another group member contributes to the project, you earn 0.4*1 = 0.4 taler.

Please be aware that the Role B participant in a group cannot contribute to the project. This participant receives the same income from the project as each Role A participant.

Step 2:

In Step 2, **only the Role B participant** in each group decides (should you have been assigned Role A, please read this part of the instructions anyway, in order to find out how a Role B participant can reach a decision). As a Role B participant, you can **reduce or leave unchanged** the income of **each** of the other participants in Step 2, namely by assigning "**points**". This becomes clear once you take a look at the input screen for Step 2:

The input screen for Step 2

Here you can see how much the individual Role A group members have contributed to the project in this period. Please bear in mind that in each period the order in which the members of your group are displayed remains the same. Group members can be identified from period to period.

Now it is up to you to decide for **each** Role A group member in this period whether you wish to allocate points and how many points you wish to distribute. Whatever you decide, you are obliged to enter a figure. If you do not wish to change the income of one particular group member, please enter 0. If you enter a number higher than 0, you reduce this group member’s income. You can move within the input fields under the heading “points” by using the tabulator key (→/) or the mouse.

If you allocate points, this costs you taler; the amount depends on the number of points you allocate. Points are **whole numbers between 0 and 20.** The more points you allocate to a member of your group, the higher your costs are. The following formula gives you the correlation between points allocated and the costs of this allocation in taler:

Cost of points allocated = Number of points allocated.

Each allocated "point" therefore costs you 1 taler. For instance, if you allocate 2 points to a member, this costs you 2 taler; if, in addition, you allocate 9 points to another group member, this costs you 9 taler; if you allocate 0 points to the two other group members, there is no cost. You have therefore allocated a total of 11 points and your **total cost** is, hence, 11 taler (2+9+0+0). If you press **Kosten-berechnung** (**Calculate cost**), the total cost is shown to you. Unless you have already clicked **Continue**, you may still change your decision.

If you choose 0 points for a particular group member, you do not change this group member's income. If, however, you allocate **one** point to a member (i.e., if you choose 1), you **reduce** this member's income by **3 taler**. If you allocate **2** points to a group member (i.e., if you choose 2), you reduce this member's income by **6 taler,** etc. **For each point that you allocate to another group member, this member's income is reduced by 3 taler.**

Please be aware that the Role A participants in a group cannot allocate points. The participant who has been assigned Role B can therefore not receive points in Step 2.

The total taler income of a Role A participant after both steps is hence calculated according to the following formula:

Taler income at the end of Step 2 for Role A = Period income for Role A = = Income from Step 1 – 3*(sum of *points* received)

The total taler income of a Role B participant is hence calculated according to the following formula after both steps:

Taler income at the end of Step 2 for Role B = Period income for Role B = = Income from Step 1 – Cost of *points* allocated by you

Please bear in mind that the taler income can also be negative for Role A participants at the end of Step 2. This could be the case whenever the income reduction from points received is higher than the income from Step 1.

Once all participants have made their decision, a screen informs you of your period income and total income thus far.

The income screen at the end of Step 2:

If you have been allocated Role A, your screen looks as follows:

Periode

1 von 10

	Sie	Mitglied 2	Mitglied 3	Mitglied 4
Gesamtsumme der Beiträge zum Projekt	[redacted]	[redacted]	[redacted]	[redacted]
Ausstattung	20	20	20	20
Beitrag zum Projekt	[redacted]	[redacted]	[redacted]	[redacted]
durchschnittlicher Beitrag zum Projekt in dieser Periode	[redacted]	[redacted]	[redacted]	[redacted]
Einkommen aus behaltener Ausstattung	[redacted]	[redacted]	[redacted]	[redacted]
Einkommen aus dem Projekt	[redacted]	[redacted]	[redacted]	[redacted]
Einkommensminderung durch erhaltene Punkte	[redacted]	[redacted]	[redacted]	[redacted]
durchschnittliche Einkommensminderung durch erhaltene Punkte in dieser Periode	[redacted]	[redacted]	[redacted]	[redacted]
Einkommen in dieser Periode	[redacted]	[redacted]	[redacted]	[redacted]
Gesamteinkommen einschließlich dieser Periode	[redacted]	[redacted]	[redacted]	[redacted]
Beiträge in allen Perioden	[redacted]	[redacted]	[redacted]	[redacted]
Einkommensminderung in allen Perioden	[redacted]	[redacted]	[redacted]	[redacted]
Beitrag in % der Ausstattung	[redacted] %	[redacted] %	[redacted] %	[redacted] %

OK

Hilfe

Sie können sich jetzt die Resultate ansehen. Drücken Sie danach bitte auf "OK".

If you have been allocated Role B, your screen looks as follows:

Your total income at the end of the experiment is the sum of the period incomes according to the following formula:

Total income (in taler) from the experiment =

= 50 + Sum of all period incomes, if the total is not negative.
Otherwise, you receive 0 taler

In addition, you are given the sum of 2.50 euro for showing up.

As mentioned above, the member of your group who has been assigned Role B will take part in the same experiment on a further occasion. At the beginning of this future experiment, the new Role A participants, who will then form a group together with your participant B, will receive a chart for their information. This chart depicts the average contributions and the individual contributions as well as the points received by the four individual Role A participants from your current group over 10 periods. The four Role A participants from the future experiment will be different participants to those in this experiment. Only the Role B participant is the same person. The participants in the new group will be told that the chart depicts the behavior of the former group with the same Role B participant.}

Do you have any further questions? If you do, please raise your hand from your booth – one of the experiment supervisors will be with you shortly.