First Impressions are
More Important than
Early Intervention

Qualifying Broken Windows
Theory in the Lab

Martin Beckenkamp
Christoph Engel
Andreas Glöckner
Bernd Irlenbusch
Heike Hennig-Schmidt
Sebastian Kube
Michael Kurschilgen
Alexander Morell
Andreas Nicklisch
Hans-Theo Normann
Emanuel Towfigh

MAX PLANCK SOCIETY

First Impressions are More Important than Early Intervention

Qualifying Broken Windows Theory in the Lab

Martin Beckenkamp / Christoph Engel / Andreas Glöckner / Bernd Irlenbusch /

Heike Hennig-Schmidt / Sebastian Kube / Michael Kurschilgen /Alexander Morell /

Andreas Nicklisch / Hans-Theo Normann / Emanuel Towfigh

June 2009

revised January 2013

First Impressions are More Important than Early Intervention Qualifying Broken Windows Theory in the Lab[*]

by

Martin Beckenkamp, Christoph Engel[**], Andreas Glöckner, Bernd Irlenbusch, Heike Hennig-Schmidt, Sebastian Kube, Michael Kurschilgen, Alexander Morell, Andreas Nicklisch, Hans-Theo Normann, Emanuel Towfigh

Abstract

Broken Windows: the metaphor has changed New York and Los Angeles. Yet it is far from un-disputed whether the broken windows policy was causal for reducing crime. In a series of lab experiments we put two components of the theory to the test. We show that first impressions and early punishment of antisocial behaviour are independently and jointly causal for cooperative-ness. The effect of good first impressions and of early vigilance cannot be explained with, but adds to, participants' initial level of benevolence. Mere impression management is not strong enough to maintain cooperation. Cooperation stabilizes if good first impressions are combined with some risk of sanctions. Yet if we control for first impressions, early vigilance only has a small effect. The effect vanishes over time.

1. Motivation

Times Square, Manhattan, 1990: clearly not the place to be. You would have met all sorts of outcasts and would have exposed yourself to a serious risk of violent crime. Times Square, Manhattan, 2000: indulge in the world's most vibrant city, at its best. Don't be afraid of violence. The crime rate is substantially below the national average.[1] Usually Mayor Rudolph W. Giuliani and New York Police Dept. Commissioner William Bratton are credited with the success (Zimring 2007). In recent years, William Bratton has repeated the New York success in Los Angeles (Wagers 2008). In both cities, he explicitly relied on the "broken windows" policy (Wilson and Kelling 1982; Skogan 1990; Kelling and Coles 1996; Sousa and Kelling 2006).

The approach was inspired by an experiment conducted by Philip Zimbardo in 1969. Zimbardo simultaneously placed two otherwise identical cars in public spaces, one in the Bronx, the other in Palo Alto. Neither car had license plates, and the hood was open. Within 26 hours the car in the Bronx was totally pillaged and destroyed, while the Palo Alto car stayed pristine for an entire week. Once the experimenters themselves broke a window with a hammer, it went to ruins within hours, even in the sheltered and prosperous Californian town (Zimbardo 1969).

Correlation analysis supports the claim that the broken windows policy, measured by the number of traffic tickets (Wilson and Boland 1978), the number of arrests per police officer for disorderly conduct or driving under influence (Sampson and Cohen 1988) or the number of misdemeanour arrests (Kelling and Sousa 2001; Corman and Mocan 2005), contributed to the decline in serious crimes, even if one controls for economic conditions and for crime deterrence (Corman and Mocan 2005) (see also Cruz Melendez 2006: for the link to the "Moving to Opportunity" Program). Along the same lines, time series evidence from Switzerland shows tougher enforcement of mild crimes to reduce the incidence of severe crimes in later years (Funk and Kugler 2003). In Los Angeles, neighbourhood deterioration preceded the onset of crime rates (Schuerman and Kobrin 1986). Yet, other studies did not find a significant effect (Novak et al. 1999; Katz et al. 2001; Geller 2007). They used a complex index of perceived social disorder as the independent variable (Sampson and Raudenbush 1999). Information about law-abiding or the number of abandoned buildings did not have a significant influence either on young males' beliefs about the risk of being convicted (Lochner 2007); (see also the mixed results by Taylor 2001; Rosenfeld et al. 2007) (further see Blumstein 1995; Bowling 1999; Messner et al. 2007: on the link to the exogenous evolution of the drug market). Yet others argue that the broken windows approach should be embedded into a broader assessment of the relationship between neighbourhood change and crime (Taub et al. 1984; Fagan 2008). Most importantly, it is far from undisputed whether correlation can be interpreted as causation (Harcourt 1998; Karmen 2000; Harcourt 2001; Sampson et al. 2002; Harcourt 2005; Harcourt and Ludwig 2006).

In this paper, we do not purport to test broken windows theory in its entirety. We are interested in two key components of the theory: (1) depending on first impressions people make in an envi-

---

[1] For details, see Uniform Crime Reports, at http://www.fbi.gov/ucr/ucr.htm.

ronment, they behave differently. Metaphorically speaking, the first broken window changes a neighbourhood. (2) If individuals quickly realize that their attempts at antisocial behaviour trigger a sanction, this tames antisocial behaviour. Specifically, we investigate whether initial experiences, both with the behaviour of peers and with their vigilance, have a beneficial effect and, if so, how long it lasts. We expect that all debating the broken windows approach would want to know whether these implications of broken windows theory hold true.

In the field, the fact that the window is not fixed (that panhandlers are free to molest passers by; that drunks congregate in the park; that rowdies menace shopkeepers) also gives a signal to those who have always been living in the area. They may read this as evidence that social cohesion is eroding. Yet normally they have many more sources of information, from which they draw their personal conclusions. They talk to each other, they read the local newspaper, they address themselves to the authorities. Therefore, in the field the effect of the signal is hard to identify (cf. Fagan 2008: 109 f. on identification problems when estimating the relationship between neighbourhood change and crime). Equally hard is identifying the motives of those who seem to behave differently. Do they move to another neighbourhood simply because they can afford it, because they want to send their children to a better school, because a new street has brought another suburb within reach – or do they move out to protect themselves from the perceived risk of crime? Is the city centre less populated because people prefer to meet in private clubs, because shopping malls in the outskirts attract customers, because people spend more time watching TV – or because they infer from the (real or metaphorical) broken windows that the centre is no longer safe?

To avoid such identification problems, in the experiments reported in this paper we create an artificial neighbourhood. The experimental setting exposes participants to a social dilemma. Individually, each participant is best off if the remaining group members contribute to a joint project while she freerides. Participants interact in a randomly composed group of four over ten announced periods. This design gives us a clean measure of (anti-)social behaviour. The less a participant contributes, the more she is selfish, and the more she imposes damage on the remaining group members.

For our first research question, the explanatory variable of interest is the impression participants happen to gather in the first period. We operationalize this as the mean contribution by the remaining three group members, in the first period. We measure the causal effect of first impressions on contributions in later rounds. First impressions do indeed have strong explanatory power. The effect does not collapse with participants' idiosyncratic social value orientation, as expressed in participants' own contribution to the public project in the first round of interaction, i.e. while they are unaware of the cooperativeness of the remaining members of their group. The average amount the remaining group members have contributed in the first round explains their choices until the penultimate round; in the final round, selfishness wins the day, even with participants who were willing to support the joint project in earlier periods. The effect of first impressions does not disappear if we control for learning, as expressed in an individual's contribution in the previous round. The effect is visible for participants who have contributed more, and for

those who have contributed less than the average of their groups in the first period. It thus is not confined to those strongly, or to those little socially minded.

Broken windows theory has been heavily used in criminal policy, as a motivation for and justification of zero tolerance with respect to petty crime. One should therefore expect that would-be offenders are more likely to desist from antisocial behaviour if they are deterred. One could further expect that community members are willing to police disorder themselves if given the opportunity, but that they are less likely to do so if they have reason to fear for revenge. This is essentially what we find. If participants are able to express disapproval and deter freeriding through costly punishment, with sufficiently favourable first impressions cooperation is stabilised in the long run, even if those punished are given a chance to strike back. If sanctions are excluded by design, cooperation decays. But conditional on first impressions, average contributions are higher, and the decay is slower.

For our second research question, the explanatory variable is reactions to antisocial behaviour in the first round of interaction. If we control for first impressions, the effect is small in early rounds, and becomes insignificant in later rounds. The critical cause is first impressions, not early vigilance. This is an important piece of news for the policy debate. In public perception, broken windows policies have been associated with being tough on crime, and on petty crime and disorder short of criminal infraction more specifically. Our data suggest that this is at most a secondary cause. If freeriders realize that crime and disorder have consequences, they behave better. This, in turn, gives others a better impression of the kind of behaviour to be accepted in this society. These impressions are key, not punishment per se.

Experiments of necessity pay a price for control. They have to abstract from many features of the real life phenomenon they aim to explain. Our experiment is no exception. We abstract from the possibility that perceived disorder attracts criminals to a community who did not inhabit it before. We are not studying the sudden change of a previously orderly neighbourhood to the worse, but have everybody start from scratch in a new environment. In our setting, disorder and crime are only distinct by the degree of antisocial behaviour, and are not qualitatively different. Loyal participants may at most fear losing some of their experimental income, not their lives, health or belongings. Despite all these simplifications, we believe the price for experimental control to be affordable.

The closest analogue in the field is the behaviour of those who newly arrive in a neighbourhood, be that a family who moves in, a child who goes to a new school, or a person who visits a new area. That way, our results also speak to the class of persons broken windows theory is most interested in: criminals who consider entering a community since, reading the signals, they believe they stand a fair chance to get away with their illegal acts.

In our experiment, there is no formal separation between disorder and crime. But through the gradual nature of our dependent variable, we have a good proxy for "criminal invasion" (Wilson and Kelling 1982): if some have been a little below others' expectations initially, chances are

others will freeride even more intensely in later periods. This is exactly how contributions decay in groups where first impressions have not been good.

In our experiments, just a few coins are at stake. In the field, the inhabitant of a neighbourhood in decay may have to leave the house in which she was born, she may see her property burglarized, and may even fear for her life. In the field, through the power of fear small initial disorder may easily start a vicious cycle. One such story might be: initial signs of disorder cause fear. Residents stay at home. This weakens social control. First offenders invade the neighbourhood. Even more residents refrain from actively maintaining order. Serious criminal activity is pulled to the neighbourhood. Yet this makes it all the more noteworthy that, in our much less dramatic setting, we also find a strong and lasting effect of first impressions.

Seemingly, the problem of criminal policy is different in that the focus is not on proactive contributions to a common good, but on the absence of antisocial behaviour. Yet as a group, the inhabitants of an area are best off if everybody's integrity and property are respected, while individually, a criminal is best off if only the others refrain from crime, and she finds ample prey. The dilemma even has a second level (cf. Yamagishi 1986; Heckathorn 1989). Individually, each member of the community is best off if others bear the cost of policing order, while she enjoys the peaceful environment. From the perspective of broken windows theory, this is not a minor issue. In their programmatic article, Wilson and Kelling claim: "The essence of the police role is to reinforce the informal control mechanisms of the community itself" (Wilson and Kelling 1982: 6).

In other respects, our experiments exactly capture the mechanism adherents of broken windows theory believe to be crucial. In our experimental groups, all rule-making is implicit and local, as are sanctions. The communities have to rely on the self policing of vague rules of conduct (Wilson and Kelling 1982). Further note that, while the theory has most frequently been used to justify rapid and strong intervention of criminal law into petty crime, per se the theory is not confined to crime. It addresses any form of socially undesirable behaviour. Therefore our testing the degree of freeriding directly fits the theory.

The remainder of this paper is organised as follows. Section 2 links our work to the related literature. Section 3 describes the dataset and the experimental designs. Section 4 presents and analyses the results. Section 5 discusses implications for broken windows theory.

2.   Related Literature

The closest analogue to our study in the legal literature is a field experiment that randomly exposed 12 of 24 matched violent crime places in Jersey City to intense police scrutiny and intervention. In the places chosen, crime rates dropped substantially, while they did not in the unaffected places (Braga et al. 1999). A further careful field experiment randomly exposed crime and disorder hot spots in Lowell, Mass. to "shallow" vs. intense police efforts to restore order, to

show that situational prevention strategies were most effective in curbing crime (Braga and Bond 2008). In a similar vein, in a series of sociological field experiments, when there were signs of disorder, like graffiti, abandoned shopping carts, or bicycles locked where they were not supposed to be, this induced passers-by also to break these and other rules (Keizer et al. 2008).

Our dataset differs from all these studies in that our "intervention" is much more light-handed; it is confined to the first impressions subjects happen to make and, if the design allows for that, to receiving punishment in the first round of interaction. Since we conducted lab experiments, we need not have second thoughts about the influence of explanatory variables beyond our control. A further advantage of our approach stems from the nature of both the dependent and the independent variables. In the field, both are categorical: people either break the law or they obey it; people either see disorder or they do not. In our setting, "disorder" is measured by the distance from socially optimal behaviour, and socially desirable behaviour is measured by the amount bystanders contribute to the joint project. Likewise, we not only observe that a participant is punished, but also how severely. We are able to distinguish between the overall level of disorder and the maximum disorder participants experience in the group of which they happen to be a member. Since all our data is from games repeated over 10 periods, we can analyse the dynamics triggered by favourable or unfavourable first impressions, and we can check when a beneficial effect of first impressions or early punishment vanishes.

In the economics literature, the closest analogue is an experiment where, in a first stage, participants were screened for their cooperativeness. In the second stage, they played a standard public-good game, knowing they were interacting with partners that scored like them in the pre-test. In a voluntary contribution mechanism, this unequivocally increased cooperation, even for those scoring low in the pre-test. However with punishment, overall contributions decayed, due to very poor performance of those scoring low in the pre-test (Gächter and Thöni 2007). The effect of sorting is positive throughout if subjects are rematched every round according to their cooperativeness in the previous round (Gunnthorsdotir et al. 2007). Likewise, if groups have a chance to exclude freeriders, this improves cooperation in a dilemma setting (Cinyabuguma et al. 2005; Croson et al. 2008), as does a mechanism that allows members to self-select into groups (Page et al. 2005), in particular if freeriders are effectively excluded by a rule that sacrifices a portion of the group income to outsiders (the Red Cross, as it was) (Brekke et al. 2009). Our study differs from this literature in that all we use is an element present in any public good game, and in any real life social dilemma: the first impressions participants happen to make, and the experience of vigilance.

Finally, we make a methodological contribution to the burgeoning field of experimental criminology (Farrington 2003; Farrington and Welsh 2005; Farrington 2006; Telep 2009). We show how meaningful and productive it is to apply standard tools from experimental economics to a longstanding issue in criminology.

3. Design and Data

In our experiments, we expose participants to a dilemma. Players interact repeatedly for 10 periods in groups of size four. The situation is fully symmetric, which all participants know. Specifically each player has the following payoff function π_i:

$$\pi_i = 20 - g_i + 0.4 * \sum_{k=1}^{4} g_k$$

Thus each period each participant receives 20 tokens from the experimenter. She is free to keep all of them, or to invest them partly or fully in the joint project. Each token she keeps gives her 1 token. Each token she invests only gives her 0.4 tokens. Yet she also receives 0.4 tokens for every token any other group member has invested into the project. Hence the entire group gains 1.6 tokens from each token invested. A participant is best off if all others have contributed fully, while she has contributed nothing. She then has 20 – 0 + 0.4*60 = 44 tokens. She is worse off if all others have contributed nothing while she alone has invested fully. She then has 20 – 20 + 0.4*20 = 8 tokens. If all contribute their entire endowments, all have 20 – 20 + 0.4*80 = 32 tokens. If all keep their entire endowments, all have 20 – 0 + 0 = 20 tokens.

In the literature, an experimental game with this structure is called a voluntary contribution mechanism (VCM). Our dataset also encompasses data from two variants. In the first variant, after all group members have decided how much to contribute to the project, they are informed about contributions by the remaining three group members. They are given the opportunity to react by spending some of their period income on reducing other group members' incomes. In the second variant, after participants have decided about punishment, players receive feedback about the punishment decision made by others and can then spend some of the remaining period income to punish those who have punished them. Since we wanted to merge our own data with data from other experimenters, we have kept the non-linear punishment technology originally used by (Fehr and Gächter 2000). It is explained in the Appendix.

Public goods experiments are a standard tool of experimental economics. In our own experiments, we moreover have used parameters that are standard in this literature. This provides us with the opportunity to test the effect of first impressions and of early vigilance in a much larger dataset. To that end, the following is partly a reanalysis of data from public good experiments that are already published (Denant-Boèment et al. 2007; Herrmann et al. 2008; Nikiforakis 2008), and partly of our own, hitherto unpublished data. The total dataset comprises 15320 datapoints, or data from 1532 participants. Table 1 informs about the different design features and parameters in more detail. All games are played in groups of four, with an endowment of 20 tokens per player. Each token contributed to the project increased each group member's payoff by 0.4 tokens.

The first column indicates whether participants had no technology for targeted sanctions (VCM), or whether they could punish each other without (Pun) or with the risk of counterpunishment (CPun). The second column indicates the origin of the data, where *MPI* denotes our own experi-

mental data, *DEN* is data provided by Denant-Boèment et al. (2007),[2] *NIK* is data taken from Nikiforakis (2008), and *HER* is data published in Herrmann et al. (2008), which consists of 16 structurally identical experiments run in different countries.[3] The third column gives the total number of individual decisions in the respective dataset. More detail on experimental procedure and on the instructions of our own, new data is to be found in the Appendix.

game-type	dataset	# obs.	P techn.	CP techn.	punishment feedback
VCM	MPI	240	-	-	-
VCM	NIK	960	-	-	-
VCM	MPI	400	-	-	-
Pun	DEN	240	FG	-	-
Pun	MPI	240	FG	-	-
Pun	NIK	480	FG	-	-
Pun	HER	10400	1:3	-	-
CPun	MPI	680	FG	FG	own
CPun	NIK	480	FG	FG	own
CPun	DEN	240	FG	FG	all
CPun	DEN	240	FG	FG	others
CPun	DEN	240	FG	FG	own
CPun	MPI	480	FG	SEV	own

Table 1
Data Structure

The fourth and fifth columns denote which punishment or, as the case may be, counter-punishment technologies were used. Here, *1:3* indicates that a linear technology was used where each punishment point assigned costs one token and reduces the other's payoff by three tokens, *FG* indicates that the non-linear technology introduced by Fehr and Gächter (2000) was used, which is described in the Appendix. *SEV* indicates that a severe technology was used, where each assigned counter-punishment point costs one token and reduces the receiver's net payoff (after the effect of received and the cost of given punishment are subtracted) by 25 %. The last column describes the amount of information that subjects were given on the counter-punishment stage, where *own* indicates that subjects only knew the amount of punishment they had received themselves, *others* indicates that subjects only knew by how much the other members of the group had been punished, and *all* indicates that subjects knew whether and by how much each subject had been punished.

---

2   The original dataset of Denant-Boèment et al. (2007) contains 20 periods. To keep datasets comparable, only the first ten periods of each matching group are considered in our analysis.

3   Athens (Number of observations N = 440), Bonn (600), Boston (560), Chengdu (960), Copenhagen (680), Dnipropetrovs'k (440), Istanbul (640), Melbourne (400), Minsk (680), Muscat (520), Nottingham (560), Riyadh (480), Samara (720), Seoul (840), St. Gallen (960), Zurich (920).

4. Results

We have two independent variables: first impressions and early vigilance. We address them in turn.

a) First Impressions

We first consider the effect of first impressions in an environment where targeted reactions to freeriding are not possible, i.e. in a voluntary contribution mechanism. Figure 1 demonstrates that the willingness to behave in a socially responsible manner strongly depends on first impressions. If the group mean was low in the first round, contributions stay very low. The higher mean contributions in the first round, the higher they are later. Eventually, contributions decay. Even excellent first impressions cannot remedy the absence of any institutional safeguard against freeriding. Yet differences in first impressions remain visible until the end of the game.
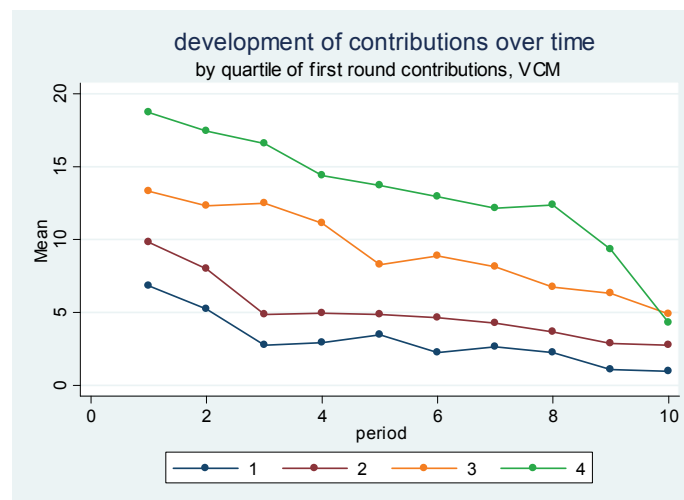
Figure 1
First Impressions in an Institution Free Environment
dv: average group contribution to public good
groups are classified by average contributions in the first round

As Table 2 shows, the visual impression is fully borne out by statistical analysis.[4] Critically, these regressions control for individuals' own contribution in the first round. It has strong explanatory power for contributions in later rounds. But even conditional on the idiosyncratic level of cooperativeness, we find a strong effect of the average contributions of the remaining group members in the first round. Actually in models 1 and 2 the latter coefficient is even larger. This

---

4   From each participant, we observe 10 contribution choices. Each individual stays a member of her group of 4 for the entire experiment. This gives us nested data. We match the data generating process by a mixed effects model. We thus estimate a random effect for groups, and another random effect for individuals nested in groups, plus residual error.

suggests that first impressions are even more important than an individual's own social value orientation. Models 2 and 3 interact first impressions with the time trend.[5]

	model 1	model 2	model 3
individual contribution in period 1	.365***	.365***	.365***
average contribution of the remaining group members in period 1	.500***	.785***	.339*
period	-.798***	-.220+	-2.247**
period^2			.169**
avf1*period		-.048***	.135**
avf1*period^2			-.015***
cons	1.398	-2.073	2.882
N	1440	1440	1400
p model	<.001	<.001	<.001

Table 2
First Impressions in an Institution Free Environment
linear mixed effects, choices nested in individuals nested in groups
data from periods 2-10
avf1: average contribution of the remaining group members in period 1
Hausman test insignificant on all models
*** $p < .001$, ** $p < .01$, * $p < .05$, + $p < .1$

One may wonder whether the significant effect of first impressions is driven by choices in early periods. Arguably, first impressions only influence first reactions, and all the rest is a result of group dynamics. Figure 2 shows that this is not the case. This figure compresses the results from nine separate regressions. From regression to regression we reduce the sample by another period, and only consider data from that period on. Except if we test the final period in isolation, for all periods do we find a strong, and a strongly significant effect of first impressions. Even if we only analyse data from a few final periods, the effect of what a participant has seen in the first period remains almost as strong as in the second period.

---

[5]   Model 3 is most informative. The model predicts that contributions decay rapidly if both this individual and the remaining group members have kept their entire endowment in the first period. The more the other group members have contributed in the first round, the more contributions are stable. The model predicts that there is no decay if the remaining group members have on average contributed 16.64 tokens (2.247/.135=16.644). Over time, the decay flattens (the quadratic time trend is positive). But this effect is most pronounced if the average contributions of others have been low in the first round (interaction between average contributions of others and the quadratic trend).
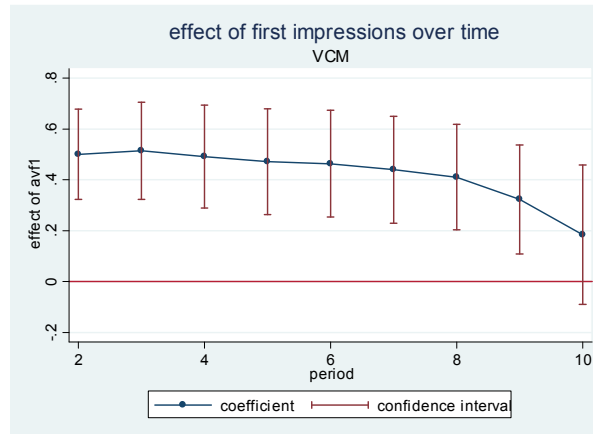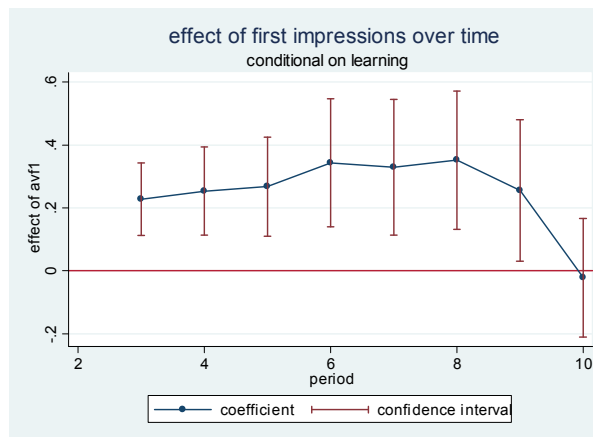
Figure 2
Long Lasting Effect of First Impressions
coefficient from model 1 of
Table 2 Table 2 for a series of regressions considering only data from the indicated period on
model for period 10 only is OLS (since the data is no longer panel data) with standard errors clustered for groups

Another competing explanation is learning. If it was true, the effect of first impressions should become insignificant once we control for a participant's contribution decision in the previous period. As Figure 3 shows, this alternative interpretation is not correct. Again with the exception of the final period, we find a significant effect of first impressions if we control for learning.[6]

Figure 3
First Impressions vs. Learning
Arrelano Bond systems estimator, one lag, robust standard errors clustered for groups
coefficient from a series of regressions considering only data from the indicated period on
model for period 10 only is OLS (since the data is no longer panel data) with standard errors clustered for groups

---

[6]   Technically, we estimate a dynamic panel with a one-period lag. Such models are known to be inconsistent, which is why we must instrument. If we use the original Arrelano-Bond estimator, the time-invariant effect of first impressions drops out. We therefore use the systems estimator, i.e. a method of moments approach. There is no mixed effects version of this estimator. We capture the dependence of observations at the group level by clustering standard errors. Again we estimate a series of eight regression (not nine regressions since the Arrelano-Bond estimator uses two lags), and reduce the sample by one period from regression to regression.

One may further wonder whether the effect of first impressions is confined to particularly selfish, or to particularly socially minded, participants, or whether, at least, it plays itself out differently for both groups. Figure 4 shows that both subgroups directly adjust to what they have seen in the first period, and then quickly converge. From period 4 on, their behaviour becomes practically undistinguishable.

Figure 4
First Impressions by Relative Position in the First Round

The regression of Table 3 supports the visual impression. Neither the two-way interaction between the average contribution of the remaining group members in period 1 and whether this participant was above or below this benchmark, nor the three-way interaction of the former with the time trend, are significant. The effect of first impressions is not different for those who are more from those who are less socially minded than their peers. The effect of first impressions does also not play itself out differently over time for these two subgroups.

average contribution of the remaining group members in period 1	.858***
above average of others in period 1	9.706***
avf1*above	-.249
period	.454*
avf1*period	-.078***
above*period	-.856**
avf1*above*period	.019
cons	-1.831
N	1440
p model	<.001

Table 3
First Impressions by Relative Position in the First Round
linear mixed effects, choices nested in individuals nested in groups
data from periods 2-10
avf1: average contribution of the remaining group members in period 1
above: dummy that is 1 if the contribution of this individual was above the average contribution of the remaining group members in period 1
Hausman test insignificant
*** $p < .001$, ** $p < .01$, * $p < .05$, + $p < .1$

12

Finally, the effect of first impressions might be conditional on the homogeneity of groups. Arguably, the overall effect might just reflect that enough groups have been sufficiently homogeneous in the first place. Table 4 shows that this is not the case. If we additionally condition choices in later rounds on the minimum contribution of one of the remaining group members, the effect of the average contribution of the remaining group members becomes even stronger.

individual contribution in period 1	.386***
average contribution of the remaining group members in period 1	.697***
minimum contribution of the remaining group members in period 1	-.161[+]
period	-.798***
cons	-.131
N	1440
p model	<.001

Table 4
First Impressions Conditional on Local Heterogeneity
linear mixed effects, choices nested in individuals nested in groups
data from periods 2-10
Hausman test insignificant
*** p < .001, ** p < .01, * p < .05, [+] p < .1

We conclude

> Result 1: In a linear public good, average contributions of the remaining group members in the first round determine contributions in later rounds. This holds irrespective of individual degree of cooperativeness. The effect lasts until the endgame effect kicks in. The effect does not collapse with learning. It is not confined to those initially above or below the average. It is not conditional on initial homogeneity.

Result 1 is important for criminal policy as a counterfactual. What is to be expected if all those who care cannot react to a perceived deterioration of socially desirable behavior, and to a vicious cycle of freeriding? Even in such an institution poor environment, first impressions have a strong, and a strongly beneficial effect. Yet they are not strong enough to stop the gradual decay of socially desirable behavior. As Figure 1 demonstrates, groups that were good in the beginning remain better than groups that had more freeriding at the outset. Whether windows are broken matters, even absent institutional intervention. Yet eventually even the best groups observe the decay of socially minded behavior. In the long run, the implicit norm is less and less obeyed.

Out there in the field, broken windows can be repaired. If someone is observed breaking a window, she may attract a reaction by bystanders. Yet intervention may be risky. Those who have pleasure from disturbing order might react aggressively against acts of vigilance. Figure 5 shows that giving those who dislike freeriding a chance to react matters strongly. This also holds if those punished have a chance to strike back. Critically for our research question, in both environments we find a strong effect of first impressions.
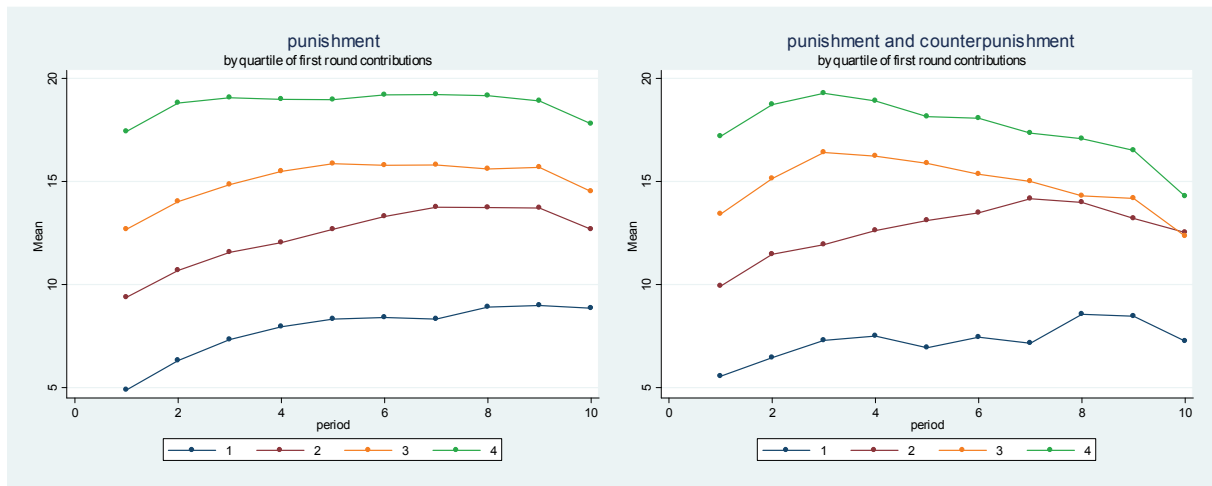
Figure 5
Effect of First Impressions in Richer Institutional Environments

Table 5 provides statistical support. In both institutional environments, first impressions again have a strong, and a highly significant beneficial effect on contributions in later rounds. Using the same tests as with a voluntary contribution mechanism, we can show that, both with punishment and with counterpunishment, the effect of first impressions is long lasting; that it does not collapse with learning; that it is not confined to particularly socially minded individuals; that it is not confined to homogeneous groups. To save space, we do not report these results in detail,[7] and confine ourselves to stating

	punishment	punishment and counterpunishment
individual contribution in period 1	.413***	.365***
average contribution of the remaining group members in period 1	.426***	.532***
period	.148***	-.169***
cons	3.559***	3.872*
N	10224	2124
p model	<.001	<.001

Table 5
Effect of First Impressions in Richer Institutional Environments

linear mixed effects, choices nested in individuals nested in groups

data from periods 2-10

Hausman test insignificant

*** p < .001, ** p < .01, * p < .05, + p < .1

Result 2: The effect of first impressions is not confined to an institution free environment. It does not disappear if those disciplining freeriders must fear revenge.

---

7   They are available from the authors upon request.

b) Early Intervention

The richer environments provide us with the possibility to test a second implication of broken windows theory. As reported in the introduction, criminal policy has mainly relied on the theory to justify zero tolerance policies. If broken windows are a problem for society, those who break them should be effectively deterred, or so the argument goes. The regressions in Table 6 and Table 7 cast doubt on this interpretation of the broken windows metaphor.[8] If punishers must not dread revenge (i.e. in treatment punishment), the main effect of the number of punishment points received in the first period on contributions in later periods is only weakly significant (p = .081) once we control for the average contribution of the remaining group members in the first period (model 2). It becomes insignificant if we interact both terms. We then only find a significant positive interaction effect. The more the remaining group members have contributed in the first round, the more effective punishment in that round is in increasing contributions in later rounds (model 3). We resurrect the main effect if we further control for the legitimacy of punishment (model 4). This we do by controlling for the fact that an individual has contributed more than the average of the remaining group members in the first period and, by the interaction term, that this person has been punished nonetheless. Such perverse or antisocial punishment happens in these experiments, and in some locations more often than in others (Herrmann, Thöni et al. 2008). But even with these additional controls, the effect of early intervention remains small.

One might object that zero tolerance policies put more stress on the certainty rather than on the severity of punishment. In a way, this is reflected in model 6. If we explain contributions in later rounds by the fact that a participant has been punished in the first round, i.e. by a dummy, then we also find a significant effect, at conventional levels, if we control for average contributions of the remaining group members. Yet once we interact both explanatory variables, in this specification neither the main effect of early punishment nor the interaction effect is significant (model 7). Even if we control for the legitimacy of punishment, we only find a weakly significant effect (model 8, p = .087).

---

8   In the typical design of a public good experiment with punishment, on which we rely both in the reanalyzed as in our own data, participants only learn whether and how intensely they have been punished themselves. We can therefore not test for a third interpretation of broken windows theory. We cannot measure the effect of *others* having been punished in the first round.

	model 1	model 2	model 3	model 4	model 5	model 6	model 7	model 8
individual contribution in period 1	.341***	.426***	.433***	.486***	.344***	.433***	.435***	.477***
amount of punishment received in period 1	.107**	.059$^+$	-.084	.085*				
punished in period 1					.809***	.583**	.424	.526$^+$
average contribution of the remaining group members in period 1		.419***	.390***	.359***		.418***	.410***	.373***
avf1*pun1			.012*					
avf1*dumpun1							.016	
contribution in period 1 above average				-.799*				-.706*
above*pun1				-.081				
above*dumpun1								-.044
period	.148***	.148***	.148***	.148***	.148***	.148***	.148***	.148***
cons	8.905***	3.373***	3.631***	3.727***	8.712***	3.119***	3.182***	3.525***
N	10224	10224	10224	10224	10224	10224	10224	10244
p model	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001

Table 6
Early Intervention vs. First Impressions: Punishment

linear mixed effects, choices nested in individuals nested in groups

data from periods 2-10

avf1: average contribution of remaining group members in period 1

pun1: amount of punishment received in period 1

dumpun1: punished in period 1

above: own contribution in period 1 above average contribution of remaining group members

Hausman tests insignificant

*** p < .001, ** p < .01, * p < .05, $^+$ p < .1

More strikingly even: once those who have been punished have a chance to strike back, whichever way we specify the regression, we never find a significant effect of early punishment (Table 7).

	model 1	model 2	model 3	model 4	model 5	model 6	model 7	model 8
individual contribution in period 1	.224***	.373***	.376***	.415***	.217***	.368***	.373***	.404***
amount of punishment received in period 1	.065	.053	-.025	.080				
punished in period 1					.116	.075	-1.258	.052
average contribution of the remaining group members in period 1		.530***	.519***	.496***		.532***	.475***	.505***
avf1*pun1			.008					
avf1*dumpun1							.113	
contribution in period 1 above average				-.488				-.539
above*pun1				-.133				
above*dumpun1								-.022
period	-.169***	-.169***	-.169***	-.169***	-.169***	-.169***	-.169***	-.169***
cons	11.737***	3.737*	3.827*	3.870*	11.847***	3.813*	4.448**	3.972*
N	2124	2124	2124	2124	2124	2124	2124	2124
p model	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001

Table 7

Early Vigilance vs. First Impressions: Punishment and Counterpunishment

linear mixed effects, choices nested in individuals nested in groups

data from periods 2-10

avf1: average contribution of remaining group members in period 1

pun1: amount of punishment received in period 1

dumpun1: punished in period 1

above: own contribution in period 1 above average contribution of remaining group members

Hausman tests insignificant

*** $p < .001$, ** $p < .01$, * $p < .05$, $^+ p < .1$

Comparing coefficients in the models of Table 6, one sees that first impressions have a much stronger effect than early vigilance. Model 5 is easiest to interpret. If a participant has been punished in the first round, however severely, she contributes a bit more than half a token more in later rounds. If the remaining group members have only contributed 2 tokens in average in the first round, this already has a stronger effect than any punishment. Figure 6 adds two more findings. Comparing the error bars, we see that, however long a panel we consider, the effect of early punishment is much more noisy than the effect of first impressions. Moreover while the effect of first impressions virtually stays identical even if we test shorter and shorter panels (confine the sample to choices from the respective period on), the effect of early punishment fades away. From period 7 on, it is no longer statistically different from zero.
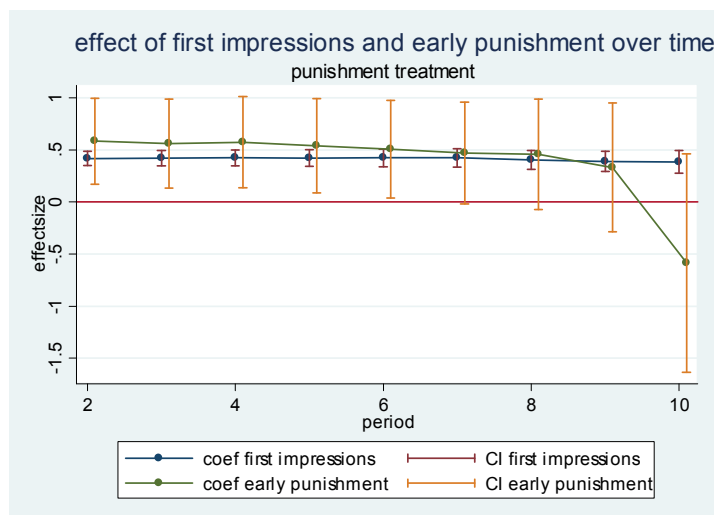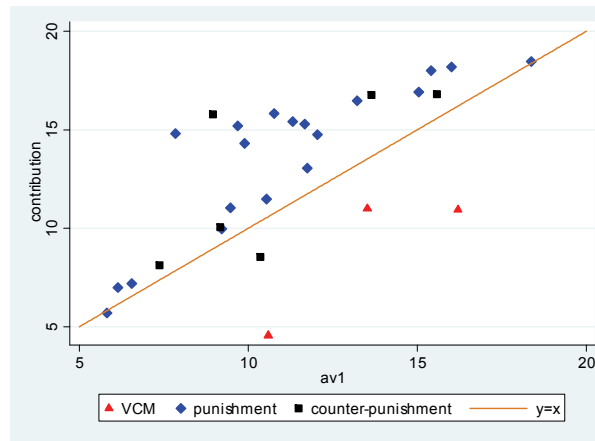
Figure 6
First Impressions and Early Vigilance in Comparison: Punishment
coefficients from model 2 of

Table 6 for a series of regressions considering only data from the indicated period on

model for period 10 only is OLS (since the data is no longer panel data) with standard errors clustered for groups

We conclude

> Result 3: First impressions are more important than early vigilance. In the long run, the effect of early vigilance fades away, while the effect of first impressions remains stable.

Figure 7 further illustrates the crucial role of first impressions. Since we have such a rich dataset, we can correlate the average contribution in the first round with the mean contribution in all later rounds. As one sees, even at this level of aggregation, results seemingly are all over the place. More disturbingly even, it seems that punishment and counterpunishment are pointless. Although in such an environment loyal participants have a chance to discipline freeriders, apparently this does not help them tame antisocial behaviour and improve cooperation. The apparent chaos dissolves once we control for first impressions. All datapoints are in the proximity of the y=x line. If there is no institution, i.e. in the VCM, they are somewhat below this line (red triangles). Otherwise they are usually somewhat above this line. The institution helps participants to even improve over the starting point.

Figure 7
First Impressions Determine Cooperation

5. Conclusions

Data from 28 experiments conducted all over the world, including five experiments run by us, demonstrate the strong effect of first impressions on cooperation in a linear public good. The average contribution of the remaining group members in the first round determines how much participants contribute to the joint project in later periods. The effect remains discernible until the end game effect kicks in. It does not disappear if one controls for learning. It is present in those who initially contributed less, and in those who contributed more than the average of the other group members. It is present in homogeneous and in heterogeneous groups, i.e. when controlling for the minimum contribution of the remaining group members in the first round. If loyal participants cannot discipline freeriders, despite favourable first impressions, contributions decay over time. If participants are allowed to punish each other, at a cost to themselves, conditional on first impressions contributions stabilize. Their level is determined by first impressions.

Early vigilance, measured by punishment received in the first round, also has a beneficial effect. Yet this effect is much smaller than the effect of first impressions. It only is present if punishers must not fear for revenge. Even absent revenge, the effect of early punishment fades away over time, while the effect of first impressions can even be found in the final period.

The closest real-world analogue to our setting is a person who is new to a neighbourhood. If this person perceives a neat environment, she expects to be treated well if she behaves well herself, and she helps maintain order if she spots signs of erosion. Note that we do not even need normativity to make this prediction. If, in addition, this person is generally willing to abide by the normative expectations prevalent in this community, of course the effect is even stronger. Neither do we need true altruists. All we need is a sufficient proportion of conditional cooperators plus, crucially, the right signals for those who newly enter the community.

In many respects, our experiments have been designed in a way that is congenial to broken windows theory. We observe the minor signs of disorder that this theory posits to be crucial. There are no explicit rules for what "order" means. Normative expectations are idiosyncratic for each context, and have to be inferred from behaviour. In other respects, we put the effect of first impressions to an even harder test: we cannot expect pre-existing social norms to guide behaviour, and there are no public officials who could help the community define expectations, and enforce them if necessary. We deprive participants of any social history, which makes the contributions of others in the first period of interaction a much noisier signal than a decay of order in a previously prosperous neighbourhood. Participants at most loose a bit of experimental money if they spot signs of antisocial behaviour, while they have reason to fear much more in the field. Therefore a vicious cycle should be much more powerful in the field.

Of course, the experimental environment is much poorer and much more artificial than a neighbourhood faced with the onset of disorder or crime. And for sure all we are testing is two components of broken windows theory: the power of first impressions and of early vigilance. Yet these limitations inherent in our method are the price we are paying for the possibility to isolate this effect, and to fully identify it.

With these obvious qualifications, our message to policymakers is straightforward. Money spent on impression management is likely to be money well spent. We can even be more specific. While good first impressions raise overall contributions in the voluntary contribution mechanism, and while they flatten the characteristic negative trend of contributions over time, they are not strong enough to reverse the trend. As many others have shown, both in the lab (Selten et al. 1997) and in the field (Ostrom 1990), for cooperation to be sustainable, vigilance and enforcement are inevitable. However, sanctions alone are also not sufficient. More importantly even for policy makers: once we control for first impressions, vigilance and sanctions at best have a minor beneficial effect. Being determined to prosecute culprits is thusnot enough. In a consequentialist perspective, it is more important to manage impressions. Beware of broken windows!

References

BLUMSTEIN, ALFRED (1995). "Youth Violence, Guns, and the Illicit Drug Industry." <u>Journal of Criminal Law and Criminology</u> 86: 531-554.

BOWLING, BENJAMIN (1999). "The Rise and Fall of New York Murder. Zero Tolerance or Crack's Decline?" <u>British Journal of Criminology</u> 86: 10-36.

BRAGA, ANTHONY A. and BRENDA J. BOND (2008). "Policing Crime and Disorder Hot Spots. A Randomized Controlled Trial." <u>Criminology</u> 46: 577-607.

BRAGA, ANTHONY A., DAVID L. WEISBURD, ELIN J. WARING, LORRAINE GREEN MAZEROLLE, WILLIAM SPELMAN and FRANCIS GAJEWSKI (1999). "Problem-Oriented Policing in Violent Crime Places. A Randomized Controlled Experiment." <u>Criminology</u> 37: 541-580.

BREKKE, KJELL ARNE, KAREN EVELYN HAUGE, JO THORI LIND and KARINE NYBORG (2009). Playing with the Good Guys: A Public Good Game with Endogenous Group Formation http://folk.uio.no/karineny/files/GoodGuys.pdf.

CINYABUGUMA, MATTHIAS, TALBOT PAGE and LOUIS PUTTERMAN (2005). "Cooperation under the Threat of Expulsion in a Public Goods Experiment." <u>Journal of Public Economics</u> 89: 1421-1435.

CORMAN, HOPE and NACI MOCAN (2005). "Carrots, Sticks, and Broken Windows." <u>Journal of Law and Economics</u> 48: 235-266.

CROSON, RACHEL T.A., ENRIQUE FATAS and TIBOR NEUGEBAUER (2008). The Effect of Excludability on Team Production http://www.economics.hawaii.edu/research/seminars/08-09/11_07_08b.pdf.

CRUZ MELENDEZ, MARIA (2006). "Moving to Opportunity & Mending Broken Windows." <u>Journal of Legislation</u> 32: 238-262.

DENANT-BOÈMENT, LAURENT, DAVID MASCLET and CHARLES NOUSSAIR (2007). "Punishment, Counter-Punishment and Sanction Enforcement in a Social Dilemma Experiment." <u>Economic Theory</u> 33: 145-167.

FAGAN, JEFFREY (2008). Crime and Neighborhood Change. <u>Understanding Crime Trends</u>. R. Rosenfeld and A. S. Goldberger. Washington, National Academy Press: 81-126.

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

FARRINGTON, DAVID P. (2006). "Key Longitudinal-Experimental Studies in Criminology." <u>Journal of Experimental Criminology</u> 2: 121-141.

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.

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

FUNK, PATRICIA and PETER KUGLER (2003). "Dynamic Interaction between Crimes." Economics Letters 79: 291-298.

GÄCHTER, SIMON and CHRISTIAN THÖNI (2007). "Social Learning and Voluntary Cooperation Among Like-Minded People." Journal of the European Economic Association 3: 303-314.

GELLER, AMANDA (2007). Neighborhood Disorder and Crime. An Analysis of Broken Windows in New York City http://ssrn.com/abstract=1079879.

GUNNTHORSDOTIR, ANNA, DANIEL HOUSER and KEVIN MCCABE (2007). "Disposition, History and Contributions in Public Goods Experiments." Journal of Economic Behavior & Organization 62: 304-315.

HARCOURT, BERNARD (1998). "Reflecting on the Subject. A Critique of the Social Influence Conception of Deterrence, tne Broken Windows Theory, and Order-Maintenance Policing New York Style." Michigan Law Review 97: 291-389.

HARCOURT, BERNARD (2001). Illusions of Order. The False Promise of Broken Windows Policing. Boston, Harvard University Press.

HARCOURT, BERNARD (2005). "Policing L.A.'s Skid Row: Crime and Real Estate Development in Downtown Los Angeles. An Experiment in Real Time." University of Chicago Legal Forum 2005: 325-404.

HARCOURT, BERNARD and JENS LUDWIG (2006). "Broken Windows: New Evidence from New York City and a Five-City Social Experiment." University of Chicago Law Review 73: 271-320.

HECKATHORN, DOUGLAS D. (1989). "Collective Action and the Second-Order Free-Rider Problem." Rationality and Society 1: 78-100.

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

KARMEN, ANDREW (2000). New York Murder Mystery. The True Story Behind the Crime Crash of the 1990s. New York, New York University Press.

KATZ, CHARLES M., VINCENT J. WEBB and DAVID R. SCHAEFER (2001). "An Assessment of the Impact of Quality-of-Life Policing on Crime and Disorder." Justice Quarterly 18: 825-876.

KEIZER, KEES, SIEGWART LINDENBERG and LINDA STEG (2008). The Spreading of Disorder. Science. 322: 1681-1685.

KELLING, GEORGE L. and CATHERINE M. COLES (1996). Fixing Broken Windows. Restoring Order and Reducing Crime in Our Communities. New York, Martin Kessler Books.

KELLING, GEORGE L. and WILLIAM H. SOUSA (2001). Do Police Matter? An Analysis of the Impact of New York City's Police Reforms. New York, Manhattan Institute.

LOCHNER, LANCE (2007). "Individual Perceptions of the Criminal Justice System." American Economic Review 97: 444-460.

MESSNER, STEVEN F., SANDRO GALEA, KENNETH TARDIFF, MELISSA TRACY, ANGELA BUCCIARELLI, TINKA MARKHAN PIPER, VICTORIA FRYE and DAVID VLAHOV (2007). "Policing, Drugs, and the Homicide Decline in New York City in the 1990s." Criminology 45: 385-414.

NIKIFORAKIS, NIKOS S. (2008). "Punishment and Counter-Punishment in Public Good Games: Can We Really Govern Ourselves?" Journal of Public Economics 92: 91-112.

NOVAK, KENNETH J., JENNIFER L. HARTMAN, ALEXANDER M. HOLSINGER and MICHAEL G. TURNER (1999). "The Effects of Aggressive Policing of Disorder on Serious Crime." Policing: An International Journal of Police Strategies & Management 22(2): 171-194.

OSTROM, ELINOR (1990). Governing the Commons. The Evolution of Institutions for Collective Action. Cambridge, New York, Cambridge University Press.

PAGE, TALBOT, LOUIS PUTTERMAN and BULENT UNEL (2005). "Voluntary Association in Public Goods Experiments. Reciprocity, Mimicry and Efficiency." Economic Journal 115: 1032-1053.

ROSENFELD, RICHARD, ROBERT FORNANGO and ANDRES RENGIFO (2007). "The Impact of Order-Maintenance Policing on New York City Homicide and Robbery Rates. 1988-2001." Criminology 45: 355-384.

SAMPSON, ROBERT J. and JACQUELINE COHEN (1988). "Deterrent Effects of the Police on Crime. A Replication and Theoretical Extension." Law and Society Review 22: 163-190.

SAMPSON, ROBERT J., JEFFREY D. MORENOFF and THOMAS GANNON-ROWLEY (2002). "Assessing 'Neighborhood Effects'. Social Processes and New Directions in Research." Annual Review of Sociology 28: 443-478.

SAMPSON, ROBERT J. and STEPHEN W. RAUDENBUSH (1999). "Systematic Social Observation of Public Spaces. A New Look at Disorder in Urban Neighbourhoods." American Journal of Sociology 105: 603-651.

SCHUERMAN, LEO and SOLOMON KOBRIN (1986). "Community Careers in Crime." Crime and Justice 8: 67-100.

SELTEN, REINHARD, MICHAEL MITZKEWITZ and GERALD R. UHLICH (1997). "Duopoly Strategies Programmed by Experienced Players." Econometrica 65: 517-555.

SKOGAN, WESLEY G. (1990). Disorder and Decline. Crime and the Spiral of Decay in American Neighborhoods. New York Toronto, Free Press ; Colleir Macmillan Canada.

SOUSA, WILLIAM H. and GEORGE L. KELLING (2006). Of "Broken Windows", Criminology, and Criminal Justice. Police Innovation. Contrasting Perspectives. D. L. Weisburd and A. A. Braga. New York, Cambridge University Press: ***.

TAUB, RICHARD P., D. GARTH TAYLOR and JAN D. DUNHAM (1984). Paths of Neighborhood Change. Race and Crime in Urban America, University of Chicago Press Chicago, IL.

TAYLOR, RALPH B. (2001). Breaking Away from Broken Windows. Baltimore Neighborhoods and the Nationwide Fight Against Crime, Grime, Fear, and Decline. Boulder, Colo., Westview Press.

TELEP, CODY W. (2009). "Citation Analysis of Randomized Experiments in Criminology and Criminal Justice. A Research Note." Journal of Experimental Criminology 5: 441-463.

WAGERS, MICHAEL LANDIS (2008). "Broken Windows Policing. The LAPD Experience." Dissertation Abstracts International Section A: Humanities and Social Sciences 68(8-A): 3603.

WILSON, JAMES Q. and BARBARA BOLAND (1978). "The Effect of the Police on Crime." Law and Society Review 12: 367-390.

WILSON, JAMES Q. and GEORGE L. KELLING (1982). "Police and Neighborhood Safety. Broken Windows." Atlantic Monthly 127: 29-38.

YAMAGISHI, TOSHIO (1986). "The Provision of a Sanctioning System as a Public Good." Journal of Personality and Social Psychology 51: 110-116.

ZIMBARDO, PHILIP (1969). "The Human Choice. Individuation, Reason, and Order versus Deindividuation, Impulse, and Chaos." Nebraska Symposium on Motivation 17: 237-307.

ZIMRING, FRANKLIN E. (2007). The Great American Crime Decline. Oxford ; New York, Oxford University Press.

Appendix: Instructions for the Severe Counterpunishment Treatment

General explanations for participants

You are taking part in an economic science experiment. If you read the following explanations closely, then you can earn a rather significant sum of money, depending on the decisions you make. It is therefore very important that you pay attention to the following points.

The instructions you have received from us are intended solely for your private information. **During the experiment, you will not be allowed to communicate with anyone.** Should you have any questions, please direct them directly to us. Not abiding by this rule will lead to exclusion from the experiment and from any payments.

In this experiment, we calculate in taler, rather than in euro. Your entire income will therefore initially be calculated in taler. The total sum of taler will later be calculated in euro as follows:

1 Taler = 4 Euro cent

In addition to the 4 euro for showing up, each participant will receive from us one instalment of **25 taler**, with which you will be able to counterbalance potential losses. **However, you will always be in a position to exclude <u>with certainty</u> the possibility of losses, with your own decisions!** The taler you will have accumulated and the 4 euro will be paid to you **in cash** at the end of the experiment.

The experiment consists of two parts. To begin with, the first part will be explained. Explanations concerning the second part will be given later.

The experiment is divided into separate periods. It consists of a total of 10 periods. Participants are randomly assigned into groups of four. Each group, thus, has three further members, apart from you. During these 10 periods, the constellation of your group of four will remain unaltered. **For 10 periods you will therefore be in the same group.** Please note that the identification number assigned to you and the other members of the group changes randomly in each period. Group members can therefore not be identified as the periods progress.

In each period, the experiment consists of **3 steps.** In Step 1, you have to decide how many taler you wish to contribute to a project. In Step 2, you are told how much all other players contributed to the project and can decide, by giving points, on whether and by how much the other group members' income from Step 1 should be increased or reduced. In Step 3, those players whose income was reduced in Step 2 can, in turn, reduce the income of the same players who did this to them.

The following pages outline the exact procedure of the experiment.

Information on the exact procedure of the experiment

<u>**Step 1**</u>

At the beginning of each period, each participant is allotted **20 taler**, which we shall henceforth refer to as his **endowment**. The player's job is now to make a decision with regard to using his endowment. You have to decide how many of the 20 taler you wish to pay into a **project** and how many you wish to keep for yourself. The consequences of your decision are explained in greater detail below.

Your **endowment** is, thus, **20 taler in each period**. You make a decision on your project contribution by typing any one whole number between 0 and 20 into the appropriate field on your screen. This field can be accessed using the mouse. As soon as you have determined your contribution, you have also decided on how many taler to keep for yourself, i.e., **20 – your contribution**. Once you have typed in your contribution, please click on **Continue**, again using the mouse. Once you have done this, your decision for this period is irreversible.

Once all members of the group have made their decisions, you will be told how high the total sum of contributions from **all** group members (including your own) to the project is. In addition, you are informed about your own contribution and the number of taler kept by you; you are also told how many taler you have earned in total during Step 1.

Your income therefore consists of two parts, namely:

(1)   the taler you have kept for yourself (**"income from taler retained"**) and

(2)   the **"income gained from the project"**. Your income from the project is .4 times the total sum of all contributions to the project.

Your total income from Step 1 is therefore calculated as follows:

Total taler income at the end of Step 1

 = income from taler retained + income from the project

Income from the project = $0,4 \times$ Total sum of all contributions to the project

The total income at the end of Step 1, in taler, is calculated according to the same formula for each member of the group.

If, for example, the sum of the contributions from all group members adds up to 60 taler, you and all other members each receive a project income of .4x 60 = 24 taler. If the group members have contributed a total of 9 taler to the project, you and all other members each receive an income of .4x9 = 3.6 taler from the project.

For each taler you keep for yourself, you earn an income of 1 taler. If, on the other hand, you contribute one taler from your endowment to your group's project instead, the sum of the contributions to the project increases by one taler and your income from the project increases by .4x1 = .4 taler. However, the income of each individual group member also increases by .4 taler, so that the group's total income increases by .4x4 = 1.6 taler. The other group members thereby also profit from your contributions from the project. In turn, you profit from other members' contributions to the project. For each taler contributed to the project by another group member, you earn .4x1 = .4 taler.

When you have finished, please click on **Continue**, using the mouse. Step 1 is now over and Step 2 about to begin.

Step 2

In Step 2, you will be told how many other group members have contributed to the project. In addition, you can **decrease, or leave as it is**, the income of **each** individual group member by giving points. All other group members are allowed to decrease **their** income, too, if they so wish.

In order to do this, you will be shown on your screen how many taler each individual group member has contributed to the project; in other words, you are told the identification number, for the current period, of each group member, as well as their contributions.

You now have to decide for **every** group member (excluding yourself) how many points you wish to give them. It is compulsory to enter a figure at this stage. If you do not wish to alter a certain group member's income, please insert 0. You can operate within the fields underneath the line **"Points"** by using the tab key (→|) or the mouse.

Periode	
1	Verbleibende Zeit [sec]: 112

Stufe 2

Gruppenmitglied	Beitrag	Punkte
Gruppenmitglied 1	XXX	
Gruppenmitglied 2	XXY	
Sie	XYX	
Gruppenmitglied 4	YXX	

Ihr Gesamtes Talereinkommen am Ende der Stufe 1: YYY

Weiter

When distributing points, you incur costs in taler which depend on the number of points you distribute to the individual players. **Distributed points are numbers between 0 and 10.** The more points you give an individual player, the higher your costs are. The total costs in taler are calculated as the sum of the costs of all points distributed to all other group members. The following table shows the connection between the points distributed to an individual group member and the costs of such distribution in taler:

Table 1: Costs of the distribution of **points** to *one* other group member in **Step 2**

Points given to a group member	0	1	2	3	4	5	6	7	8	9	10
Cost of these points in taler	0	1	2	4	6	9	12	16	20	25	30

Your total cost of the points distribution is the sum of all costs to all three other group members. For example, if you have allocated 2 points to one member, your cost is 2 taler; if, in addition, you give 9 points to another group member, your cost is 25 taler; if you give the final group member 0 points, you have no costs. The **total cost** to you is therefore 27 taler (2+25+0). As long as you have not yet clicked on **Continue**, you may still change your decision.

If you choose 0 points for a certain group member, you do not alter this group member's income. With each **point** allocated to a group member, you **decrease** this particular group member's taler income from Step 1 by **10 per cent**. Thus, if you allocate **2 points** to a group member, for instance, thereby choosing 2, you decrease his income by **20 per cent.** The points allocated by you therefore determine how significantly one group member's taler income from Step 1 is reduced.

Whether, or by how much, a group member's income from Step 1 is reduced **overall** depends on the total number of points received. If, for instance, one member receives **a total of 3 points** from all other members, the income in Step 1 is reduced by **30 per cent**. If a member receives a total of **4 points**, the income in Step 1 is reduced by **40 per cent**. If a member receives exactly 10 points or more, the income in Step 1 is reduced by 100 per cent. The income in Step 1, in this case, would be reduced to Zero for this member. Your total income from the first two steps, in taler, is thus calculated as follows:

Total taler income at the end of Step 2:

$=$ (Total taler income after Step 1) $\times$ (10 – points received)/10
– cost of points distributed by you

if points received < 10

$= -$ cost of points distributed by you

if points received ≥ 10

Step 3

In the third and final step, you are told how many points each individual group member has given you. **If group members have given you points in Step 2, you can now reduce the income of these group members by allocating what is known as "*counter-points*".** Only those group members who received points in Step 2 are allowed to allocate counter-points. And these counter-points can only be distributed to group members who gave them points in Step 2.

A counter-point reduces the income that remained in the possession of the member in question at the end of Step 2 **by 25 %.** Should a member receive exactly 4 or more counter-points, the income from Step 2 is reduced by 100%. If you yourself receive 4 or more counter-points from group members to whom you gave points in the previous step, your own income from Step 2 is therefore also reduced by 100%.

The costs of counter-points are calculated just as in Step 2. Note, however, that if you give one group member counter-points in addition to having given him points, then the costs are calculated according to the sum of all the points this group member has received from you in Steps 2 and 3.

The costs of the counter-points can be seen in Table 2. **Example:** If you give Player 1 a total of 2 points in Step 2, your cost in Step 2 is 2 taler. If you give Player 1 a total of 3 further points in Step 3, a further 7 taler are added to your cost.

Table 2: Costs of the distribution of **Counter-points** to *one* other group member in **Step 3**

Points you have already given to the group member in Step 2	Counter-points given to the group member in Step 3 Gruppenmitglied in Stufe 3				
	0	**1**	**2**	**3**	**4**
0	0	1	2	4	6
1	0	1	3	5	8
2	0	2	4	7	10
3	0	2	5	8	12
4	0	3	6	10	14
5	0	3	7	11	16
6	0	4	8	13	18
7	0	4	9	14	
8	0	5	10		
9	0	5			
10	0				

On your screen, you can see how many points each individual group member has given to you in Step 2. Now you must decide, for **each** of these group members, how many counter-points you wish to give this member. It is compulsory to enter a figure at this stage. If you do not wish to alter a certain group member's income, please insert 0.

Your total income from all three steps, in taler, is thus calculated as follows:

Total taler income at the end of Step 3 = Period Income

= (Total taler income after Step 2) × (4 – counter-points received)/4

– cost of counter-points distributed by you

if the sum of the counter-points received is < 4

= – cost of counter-points distributed by you

if the sum of the counter-points received is ≥ 4

You will also find this information on the final screen of each period.

Periode				Verbleibende Zeit [sec]: 106
	1			

Stufe 3

Gruppenmitglied	Punkte erhalten	Gegen-Punkte
Gruppenmitglied 1	xxx	
Gruppenmitglied 2	0	
Sie		
Gruppenmitglied 4	yxx	

Ihr gesamtes Talereinkommen am Ende der Stufe 2: yyy

Weiter

The Payoff

Your total income, in taler, is calculated from the sum of your taler income in each period, in addition to the flat payment of 25 taler given to you at the beginning. As mentioned above, you receive 4 euro cent for each taler. You are also paid 4 Euro for showing up.

Do you have any further questions?