Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

FAKULTÄT
FÜR WIRTSCHAFTS- UND
SOZIALWISSENSCHAFTEN

What goes around, comes around: Experimental evidence on exposed lies

Kai-Uwe Schnapp
Sarah Mörtenhuber
Andreas Nicklisch

WiSo-HH Working Paper Series
Working Paper No. 16
June 2014

WiSo-HH Working Paper Series
Working Paper No. 16
June 2014

What goes around, comes around: Experimental evidence on exposed lies

Kai-Uwe Schnapp, Institute of Political Science, Universität Hamburg
Sarah Mörtenhuber, School of Economics and Social Sciences, Universität Hamburg
Andreas Nicklisch, Department of Economics, Universität Hamburg, and Max Planck Institute for Research on Collective Goods

ISSN 2196-8128

Font used: „TheSans UHH" / LucasFonts

Die Working Paper Series bieten Forscherinnen und Forschern, die an Projekten in Federführung oder mit der Beteiligung der Fakultät für Wirtschafts- und Sozialwissenschaften der Uni-versität Hamburg tätig sind, die Möglichkeit zur digitalen Publikation ihrer Forschungser-gebnisse. Die Reihe erscheint in unregelmäßiger Reihenfolge.

Jede Nummer erscheint in digitaler Version unter
https://www.wiso.uni-hamburg.de/de/forschung/working-paper-series/

Kontakt:
WiSo-Forschungslabor
Von-Melle-Park 5
20146 Hamburg
E-Mail: experiments@wiso.uni-hamburg.de
Web: http://www.wiso.uni-hamburg.de/forschung/forschungslabor/home/

What goes around, comes around: Experimental evidence on exposed lies*

Kai-Uwe Schnapp[§] Sarah Mörtenhuber[‡] Andreas Nicklisch[¶]

June 17, 2014

Abstract

We investigate experimentally the optimal way to handle the exposure of a noble lie. Specifically, we consider the provision of a public good where the feedback for players is manipulated in order to foster contributions. Then, we reveal the feedback manipulation and analyze whether the public good is provided more efficiently if the same feedback manipulation is applied again, or if truthful feedback is provided. We find that continuing with the manipulated feedback initially leads to significantly lower contribution rates, while the difference between treatments disappears over time. On the other hand, receiving honest feedback harms cooperation within groups more substantially than a potential noble lie if groups are very heterogeneous in terms of their contributions. Therefore, we cannot provide a clear-cut answer whether or not one of the two feedback mechanisms is more beneficial for society.

Keywords: *Feedback manipulation, Noble lie, Public goods provision, Truth*
JEL-Classification: C91, D03, H41

---

*We thank Thomas Lauer and Bettina Rockenbach for providing detailed information on their experiment. Financial support by the University of Hamburg is gratefully acknowledged.

§ *Corresponding author:* Institute of Political Science, University of Hamburg, Allendeplatz 1, 20146 Hamburg, Germany, e-mail: kai-uwe.schnapp@wiso.uni-hamburg.de

‡ School of Economics and Social Science, University of Hamburg, von-Melle-Park 5, 20146 Hamburg, Germany, e-mail: sarah.moertenhuber@wiso.uni-hamburg.de

¶ Department of Economics, University of Hamburg, and Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Allee 10, 53110 Bonn, Germany, e-mail: andreas.nicklisch@wiso.uni-hamburg.de

1 Introduction

It is common wisdom that lying is endemic to politics. Most citizens of any state claim that they experience some political lies at all levels of the government. A specific problem in the debate are politicians lying in democratic systems, and there are various positions about its legitimacy. Some say that a political lie can help to actually reach a public good everybody wants (Dixon, 2002) and is therefore justifiable. Therefore, supporters of this approach refer to political lies seeming to be "the right thing to do" at the time of lying (with reference to Plato) as royal or noble lies. On the other hand, some scholars state that lying is so foreign to democracy that it does nothing but undermine democracy itself (Aughey, 2002), even if it produces welfare benefits. Stressing the moral dimension of lying, opponents of the noble lie claim that "a particular act of government may be exactly the right thing to do in utilitarian terms and yet leave the man who does it guilty of a moral wrong" (Walzer, 1973). Thus, it does not have any place in democratic politics whatsoever.

Following up the discussion concerning the intermediate consequences of lying for level of cooperation of society members within democracies – and being somewhat agnostic whether a noble lie is morally appropriate or not – our study attempts to shed light on the aftermath of royal lies. That is, in our paper we analyze the question whether there is an optimal way for politicians to proceed if a noble lie has been exposed. In other words: once the noble lies have been discovered, we ask whether it is more beneficial for society that politicians continue to tell reassuring lies or to provide the (perhaps) inconvenient truth.

There are two potential answers to this question. First, one could claim that facts about betrayals for the sake of society are worse than concerns over potential betrayals: if subjects are aware of having been betrayed and the extent to which their goodwill is abused, the consequences in terms of cooperativeness are worse than the consequences of an uncertain, potential betrayal. Cases of potential betrayals are likely not to negatively influence the cooperation of players as they cannot be identified as betrayals with certainty. On the other hand, one could argue that concerns over potential betrayals are worse than facts about betrayals. That is, if subjects are aware of the extent to which their goodwill is abused, the consequences in terms of cooperativeness are less severe than the consequences of an uncertain potential betrayal.

With these two claims in mind, we consider a series of experimental public goods game. Every participant in the experiments faces two sequences of 20 periods of a voluntary contribution mechanism game (VCM) each. In all treatment conditions, feedback to the players in the first sequence is manipulated such that the degree of cooperation within the experimental group has been increased. That is, the group of players as a whole gains at the cost of some very cooperative players and by the help of a noble lie. Between the two sequences the feedback manipulation is exposed to participants. Thus, the confidence in the current feedback mechanism is sufficiently shaken. In one treatment condition players then play the second sequence of VCM with the same feedback mechanism. In a sense, the noble lie is continued in this treatment condition. In a second treatment condition we introduce a non-manipulated feedback mechanism and inform players on the new nature of the feedback. We compare cooperation rates between the two treatment conditions.

As such, our experiment extends a recent experimental study by Hoffmann et al. (2013). The authors analyze different types of royal lies in the same public good setting.[1] Particularly, they are interested in the question which type of royal lie leads to superior outcomes, even though the confidence is sufficiently shaken by revealing the lie's nature. Comparing the cooperation rates under different types of royal lies and under non-manipulated feedback, Hoffmann et al. find that general royal lies do not travel far, while lying selectively to very cooperative players in order to keep them going sustains and improves the social welfare significantly. More specifically, in the selective lying mechanism, subjects receive feedback on the average cooperation rate of others in their respective group. However, whereas subjects who cooperate less than or exactly the average to the public good are shown the actual average, subjects who gave more than average in the previous round are shown their own cooperation rate *as* the average.

To understand Hoffmann et al.'s findings, they have to be related to earlier findings on the motivation of players to cooperate in the voluntary contribution

[1]Earlier studies by Marwell and Ames (1981) and Weimann (1994) deal also with manipulated feedback and find no significant effect. However, both studies provided feedback in a different format (i.e., percentage of the total contributions) which may not trigger the same response.

mechanism game. A large body of evidence shows that the majority of participants in this game cooperates conditionally (e.g., Fischbacher et al., 2001, Kocher et al., 2008, Fischbacher & Gächter, 2010). That is, participants try to match the average cooperation level of all other players in the group:[2] Participants are very cooperative in very cooperative groups, while they defect if all other players do so as well. One constellation that players attempt to avoid particularly is that they cooperate while all other group members do not do so ("sucker aversion", cf. Fischbacher & Gächter, 2010).

This is where feedback about others' contributions comes in. To sustain cooperation, it seems to be important to avoid that cooperators feel like suckers,[3] and this is exactly what Hoffmann et al.'s selective lying does. Based on this result, we analyze how different feedback mechanisms and different levels of uncertainty influence subjects' willingness to cooperate after the lie has been exposed. Our findings provide a number of important insights: first, in the initial periods of the second sequence, average cooperation with truthful feedback tends to exceed (weakly significant) the average cooperation in the treatment condition with the noble lies. However, as the second sequence of the experiment proceeds, differences between treatment conditions vanish. Thus, what goes around, comes around: both, the reassuring lie which potentially hides betrayal and the inconvenient truth about it eventually leads to a break down of cooperation within groups. Yet, they do so at different speed: in more homogeneous groups (in terms of their degree of cooperation) the reduction in the cooperation rate is larger after seeing a potential noble lie than after being accurately informed about the extent to which other players free-ride on one's cooperation. However, the opposite is true for very heterogeneous cooperation rates within groups. In other words, if the inconvenient truth about the betrayal is overwhelmingly bad, the

[2]Neugebauer et al. (2009) and Fischbacher and Gächter (2010) show that participants actually try not to match the average, but to cooperate slightly less than the group does on average.

[3]Related to this issue, Pillutla and Chen (1999) and Nikiforakis (2010) analyze whether cooperative feedback (contributions feedback) or competitive feedback (earnings feedback) establishes a higher degree of the overall cooperation within the group. Results partially show that the first feedback format induces more pro-social behavior and increases the overall cooperation rate significantly. Irlenbusch and Rilke (2013) test the effect of selective feedback information in public good games. Only if the selection criteria for feedback are unknown to the participants, samples of very cooperative players boost cooperation while samples of very uncooperative players hamper them.

consequences in terms of cooperation are worse than the consequences of a reassuring information, despite the fact that most likely players assume that the information is wrong. In turn, in rather homogeneous groups the latter consequences are worse than the former. Finally, a third treatment condition which we introduce in order to disentangle the effect of uncertainty about being betrayed and uncertainty about the betrayal's dimension does not provide clear answers. Therefore, it seems that both types of uncertainty matter and eventually lead to a decline in cooperation rates.

Our paper is organized as follows: We first discuss the implications of lying for interactions within democracies from a political science point of view. Then we explain the structure of the game we have our subjects play as well as the experimental procedure. In the following section we describe and analyze the results of the experiment. We conclude with an answer to our questions and an outlook on future research concerning lying in politics.

2 Democracy and Lying

Let us start with a general definition of lies. According to Mahon (2008), lying means that people make a believed-false statement to another person with the intention that this other person believes the statement is true. For our purpose, it is useful to differentiate between private and public lies. A public lie is being told by people or groups who hold any kind of public office in the execution of this office. Private lies, on the other hand, are all lies that are used in private relations between people. Despite the broad academic attention towards private lying, we focus on public lying and leave major parts of the literature on private lying aside.

The public lie can be differentiated into a selfish lie and the "noble lie". While it is self-evident, that the selfish lie is morally bad, this is not so clear for the noble lie, that is, a public lie may enhance some kind of public good. Some people say that a noble lie may be justifiable, even in the context of a democratic polity. It should be used if it is really necessary and if it has a strong justification only. However, if circumstances really demand a noble lie, the noble lie may be told. The contrary position holds that a noble lie cannot be justified, at least not in a democratic polity. Proponents of this position argue that this is not only a question of individual morale, but a question that touches the very fundamentals of the idea of democracy. Democracy relies on the participation of those who are affected by a decision in this decision (Dahl, 1998). This participation requires accurately informed citizens for effective decisions. Lying obviously contradicts this principle. People who are being lied to, and be it for the better of the public good, cannot be described as well informed. Sissela Bok puts it slightly differently:

> "Deception ... strikes at the very essence of democratic government. It allows those in power to override or nullify the right vested in the people to cast an informed vote in critical elections. Deceiving the people for the sake of the people is a self-contradictory notion in a democracy" (Bok, 2004, 375–376).

There is yet another more pragmatic argument against democratic lying. Democratic governments do heavily rely on their citizens' trust if they want to govern successfully. With public lying, even in its noble disguise, citizens loose this trust and therefore governments loose the basis for their political success. This means that in the long run public lying will always be detrimental to the well-functioning of democracy and its successful provision of public goods.

Dixon (2002) discusses pros and cons of the democratic lie against the backdrop of a real example: the political developments leading to the Good Friday agreement on a cease-fire in Northern Ireland. The public good involved in this case was peace in Northern Ireland, something so many people longed for badly. However, in order to achieve this goal people on both sides of the conflict needed to be told (by their elites) that the other side was (already) delivering to justify that it was time for them to deliver as well. Deliverance occurred, since every side believed that the other was already delivering (which, at least at the elite level, was true). According to Dixon, a problem may result from the cheating of the elites themselves and from the possible self-defeating effects of this cheating, once the political trick was exposed by some news media or other source.

As a consequence, Dixon discusses three possible positions towards political lying. (1) The absolutist position of course holds that lying is bad (Dixon, 2002,

737) and does not work in a democracy, since it undermines its very principles. (2) The realist position holds that the dishonesty may be the smaller of two evils, because it may lead to the provision of a public good that could not be obtained otherwise (the Northern Ireland peace agreement is assumed to be a good example). (3) The democratic realist position than takes some middle ground: morale leadership requires sometimes doing something wrong. Hence, the boundary between persuasion and manipulation is fuzzy at times (Aughey, 2002, 5). Yet, as time proceeds, the gap between political spin and backstage political reality has to be as narrow as possible (Aughey 2002, 5).

Aughey takes Dixon's arguments literally. He asks two questions: Can lying be morally appropriate in a democracy? And if so: Can it be effective? His answer is rather complex. Nobility of a lie is supposed to be in its intent and in its consequences (Aughey 2002, 6). The difficulty for democracy is that the "noble lie" presumes differences between politicians and citizens in their knowledge and understanding of political processes (Aughey 2002, 6). However, this presumption does not correspond to the normative basis of democracy, as discussed above. In addition, Aughey argues that a "right to lie" on the politicians' side implies a "right to suspect" on the citizens' side. Of course, the question is whether suspicion is a sound basis for democratic politics. This question leads us to the core of our experiment: Does political lying have consequences for (democratic) politics in the long run and if it does, what are these consequences?

We demonstrate that the reservations against noble lies, at least after their exposition, are correct. Lying in public realm means that in the long run citizens cannot distinguish any longer between truth and lie. Citizens do not know what to believe and what not. "Reasoned political communication ... becomes impossible" (Aughey 2002, 15) and, as a consequence, trustworthy cooperation becomes unlikely. Political lying, even if equipped with noble intentions, undermines the very basis of democracy because it eliminates a trustful relationship between citizens and those in the government office.

3 The Game

For the analysis of our research question we use the linear standard VCM (e.g., Isaac et al., 1985) as an experimental tool. This design has widely been tested experimentally by biologists, economists and political scientists (see Zelmer, 2003, Chaudhuri, 2011, for surveys). The VCM allows us to investigate cooperation behavior in response to different feedback regimes. It is not tailored to a specific case or application in international politics, but serves as a framework that incorporates important features of many relevant situations.

The game is repeatedly played in groups of four participants. The composition of the group remains unchanged over the entire duration of the experiment. In total, participants played two sequences of 20 periods of the following stage game each: at the beginning of each period, each player receives an endowment of 20 ECU (experimental currency units). Players simultaneously decide how many ECU should be contributed to the public good, g_i, with $g_i \in \{0, 1, 2, ..., 20\}$. Each ECU contributed to the public good yields a benefit of 0.4 ECU (the marginal per capita return) to every player in the group (including i). Each ECU not contributed to the private good is privately kept and yields a benefit of 1 ECU to player i only. Therefore, player i's individual payoff function is $\pi_i = 20 - g_i + 0.4 \sum_{j=1}^{4} g_j$.

All parameters of the game are commonly known by all participants of the experiment. Since subjects lose 0.6 ECU per ECU contributed (recall that the marginal per capita return is 0.4), the individually optimal strategy of the game is to keep the whole endowment of 20 ECU for themselves. In turn, since 4 times 0.4 equals 1.6, the group as a whole benefits from each ECU contributed, such that the social optimal strategy of the game is to contribute the entire endowment of 20 ECU to the public good. The rationale of individual and social optimal strategies remains unchanged if the game is played repeatedly over a finite number of periods: since everybody knows that the individual optimal strategy is to keep every ECU in the game's last period, a deviation from this strategy in the second last period is useless, therefore a deviation from this strategy in the third last period is useless, and so forth. In other words: individual contributions to the public good reflect the voluntary cooperation rate of players.

In order to analyze the response to exposed lies, we divide the game into two separate parts. In sequence one, we implement the partially manipulated feedback mechanism, first applied by Hoffmann et al. (2013). Here, participants play 20 periods of the VCM and receive at the end of each period feedback on the individual contributions as well as on the average contribution of the other three members of the group (rounded to the next full integer). However, subjects contributing more than the average are shown their own contribution of this period, while all other subjects receive an accurate feedback on the average contributions of all other players. Subjects are informed about possible deviation of the feedback from the actual value – but not about the exact feedback mechanism – in the instructions prior to the experiment.[4]

It follows for players contributing more than the average that their impression concerning the cooperativeness within the group is too optimistic. Consequently, greater optimism concerning the average contributions of other group members leads to an increase or at least a preservation of current contribution rates. In turn, this has additional positive implications with respect to the cooperativeness of the other group members: since they face a high average contribution rate by the other group members (and receive correct information on this), they increase or preserve at least their current contribution rates as well. Overall, both correctly informed and deceived subjects benefit in terms of overall cooperativeness within their group (Hoffmann et al., 2013).

Prior to the start of sequence two, we reveal the exact feedback mechanism to all participants. In other words, participants are informed that they receive correct signals once they contribute at most the average, and that they receive their own contribution as signals otherwise. Sequence two implements another 20 periods of the VCM with feedback information that differs between three treatment conditions:

(1) In the LIE condition, again the partially manipulated feedback mechanism is implemented. That is, players' confidence in the accuracy of the feedback mechanism is substantially unsettled. In this treatment condition, receiving the feedback that the average of the others' contributions equals their own contribution causes two types of uncertainty: on the one hand, they may have contributed

[4]This is important as to honor the convention of experimental economics not to lie to subjects.

more than the average (but they may have coordinated alternatively on the average contribution). On the other hand, given false feedback, they do not know the degree of betrayal (i.e., the amount by which their own contribution surpasses the average).[5]

(2) In the WARNING condition, we eliminate one of the two types of uncertainty: here, the partially manipulated feedback mechanism is implemented. However, whenever players receive a potentially wrong signal (i.e., the feedback that the others' average contribution equals the own contribution), they receive an additional signal stating that this information is wrong if the signal is actually wrong. Therefore, players contributing more than the average in this treatment condition know that this is the case, while they also face the uncertainty concerning the degree of betrayal.

(3) Finally, in the TRUTH condition, we eliminate both types of uncertainty. In other words, all participants receive accurate signals concerning the others' average contribution in this treatment condition.[6]

Certainly, revealing the exact feedback mechanism between the two parts shakes players' confidence concerning the cooperation within the group substantially. Considering players' common inclination to conditionally cooperate, contribution rates decrease from sequence one to sequence two in all treatment conditions. The crucial questions is which way to deal with the betrayal in sequence two is the least harmful for contributions to the VCM. We test the three treatment conditions regarding their influence on the degree of cooperativeness first of all for possibly betrayed players, and moreover, due to the feedback mechanism of conditional cooperation, for the entire group. LIE causes a high degree of uncertainty concerning whether or not there is a betrayal of high contributors and the extent to which those players are betrayed, TRUTH provides certainty on both the betrayal and its extent, while WARNING is somewhat between to other two treatment conditions. Therefore, following the claim that facts about be-

[5] At this point, the advantage of partially manipulated feedback mechanisms becomes clear: revealing a uniformly manipulated feedback mechanism (e.g., adding ten percent to every contribution for the signal) would lead to no uncertainty afterwards once we reveal the mechanism – players would simply recalculate the signal, while they cannot do so under partially manipulated feedback.

[6] We abstain from showing the players a wrong signal along with the warning that this information is wrong and the accurate signal, but show the accurate signal only.

trayals are worse than concerns over potential betrayals, we hypothesize that the decrease in contributions between sequence one and two is the largest in TRUTH, while it is smaller in WARNING, and smallest in LIE. On the other hand, one may follow the claim that concerns over potential betrayals are worse than facts about betrayals. In this case, the reverse hypothesis follows: the decrease between parts is the largest in LIE, while it is smaller in WARNING, and smallest in TRUTH.

An experimental session proceeds the following way: once all subjects were seated, written instructions on sequence one were handed to them before the experimenter read them out aloud.[7] Additionally, they were given the opportunity to ask questions in private. Before the experiment started, subjects had to answer a set of control questions. Subjects were not informed about the content of sequence two prior or during sequence one of the experiment; subjects received a new set of instructions at the start of sequence two. At the end of the entire experiment, subjects had to answer a short socio-demographic questionnaire along an unincentivized hypothetical statement on their preference for or against their feedback mechanism in comparison to a truthful mechanism.

Overall, we ran nine sessions with a total of 176 participants who were matched randomly into groups of four. For one group being an independent observation, we collected 16 observations in LIE, 11 in WARNING and 17 in TRUTH. Each subject participated in only one treatment condition. The experiment was run in June and July 2013 in the experimental laboratory of the Faculty of Economic and Social Sciences at the University of Hamburg, Germany. The experiment was programmed using the software package zTree (Fischbacher, 2007) and hroot (Bock et al., 2012) was used for recruiting. Most participants were students (2% non-students) with different academic backgrounds including economics, 56% were women and median age was 24. Sessions lasted typically 60 minutes and the average payoff was 11.43 Euro (subjects were paid in private immediately after the end of a session).

[7]English translations of the experimental instructions are enclosed in the appendix.

4 Results

In this chapter, we will present the results of the experiment. In a first step, we discuss overall differences between treatment conditions. In a second step, we compare individual differences between responses to false feedback information in LIE and WARNING to responses to correct information that players receive in TRUTH. Finally, we assess the welfare consequences of the different feedback mechanisms.

Let us start with the average contributions in the first and the second sequence of the experiment. Figure 1 shows the development of average contributions by periods and across treatment conditions. As expected, the development within the first sequence remains very similar across treatments. It furthermore is in contrast to the typical evolution of voluntary cooperation in a public good game where contributions decrease constantly over time (cf., e.g., Chaudhuri, 2011), contributions remain stable at a level of about 11 out of 20 ECU (in particular, mean contributions are 11.2 in LIE, 11.7 in WARNING, and 11 in TRUTH). Comparing the contributions between treatment conditions, both for the first period and across all periods of the first sequence, do not reveal any significant differences.[8]

The picture changes in the second sequence of the experiment. After the revelation of the feedback mechanism, players' confidence in the cooperation rate within groups is shaken and contributions constantly decrease in the course of the second sequence in all treatment conditions.[9] Yet, it seems that they do so at different speed: There is a substantial restart effect in WARNING and TRUTH in the sense that average contributions start in period 21 in WARNING at 10.5 and in TRUTH at 11, whereas mean contributions in LIE drop to 8.7. Notice that

[8] $p > .5$ for all comparisons; we test group averages using two-sided Wilcoxon Mann-Whitney rank sum tests.

[9] Our results do not replicate the findings of Hoffmann et al. in the second sequence of the experiment. Their data suggest that contributions remain somewhat stable, even after the lie has been exposed. Since both studies apply different exposure mechanisms for the lie between the first and the second sequence, it seems very likely that this mechanism influences crucially the further development of cooperation rates in the experiment: Hoffmann et al. inform players that they see "at least their own contribution as the group average," while we state that they receive correct signals once they contribute at most the average, and that they receive their own contribution as signals otherwise.

the difference between group averages in TRUTH and LIE is weakly significant.[10] Similarly, the differences between group averages in TRUTH and LIE in the next three periods are (almost) weakly significant.[11] Thus, there is no evidence that LIE enhances the cooperation in the second sequence of the experiment. Rather, there is some evidence that LIE yields an overall less cooperative start in terms of contributions than TRUTH in the initial periods of sequence two. WARNING seems to be located somewhat in-between the other two treatment conditions. However, as the second sequence of the experiment proceeds, treatment differences disappear.

Result 1. *Average contributions under* TRUTH *tend to exceed (weakly significant) average contributions under* LIE *in the initial periods of sequence two. However, treatment differences disappear during the course of the experiment.*

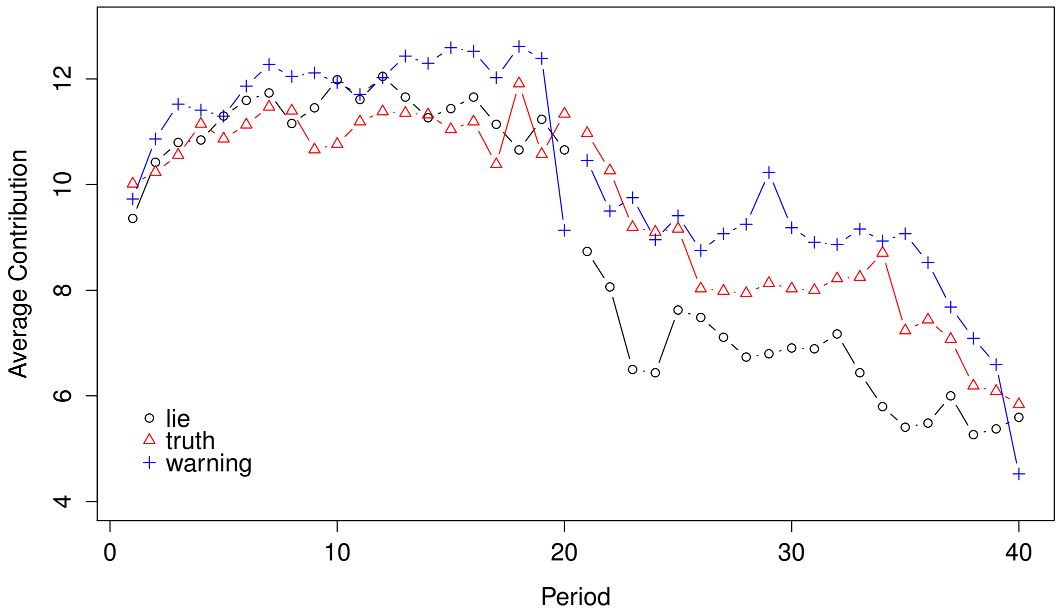

Figure 1: Average contributions by periods across treatments

To gain more insight on the dynamics in the different treatment conditions, we analyze individual responses to received feedback in a series of individual random

[10] $p = .08$ using two-sided Wilcoxon Mann-Whitney rank sum tests.

[11] $p = .15$ in period 22, $p = .11$ in period 23, and $p = .06$ in period 24, again using two-sided Wilcoxon Mann-Whitney rank sum tests.

effect regressions. More specifically, we estimate an ordinary least square regression on the first differences of contributions in the second sequence of the experiment. That is, our dependent variable is $\Delta_i := g_i^{t+1} - g_i^t$ with $t \in \{21, 22, ..., 39\}$, while we estimate regressions for each treatment condition separately. We consider as independent variables the period t as cooperation rates decline towards the end of the experiment. In additions, we include variable dif_-^t which we define as $dif_-^t = \max\{0, \sum_{j\setminus i} g_j^t/3 - g_i^t\}$. That is, dif_-^t measures the difference between the contribution of player i and the average contributions by other group members if i contributes less than the average. This information is accurately provided to all players in all treatment conditions. Likewise, in TRUTH, we include the corresponding variable $dif_+^t = \max\{0, g_i^t - \sum_{j\setminus i} g_j^t/3\}$ which measures by which i's contribution exceeds the other players' average contribution if i contributes more than the average. Notice that we cannot include dif_+^t into the regressions for LIE and WARNING as players see ambiguous feedback in this case, that is, their own contribution as the average. Therefore, in the two corresponding estimations, the constant term reveals i's response in terms of Δ_i to this feedback information. Estimation results are summarized in Table 1. We report estimated coefficients along robust standard errors. The number of observations (*obs*), independent observations (*nobs*) are reported; asterisks indicate significance levels. The fitness of the models is tested on the basis of Wald-Chi^2-tests.

	TRUTH	LIE	WARNING
period	0.006 (0.015)	0.040*** (0.013)	−0.053** (0.022)
dif_-^t	0.299*** (0.081)	0.487*** (0.090)	0.383*** (0.134)
dif_+^t	−0.517*** (0.112)		
constant	−0.088 (0.521)	−2.362*** (0.467)	0.515 (0.644)
obs	1292	1216	836
nobs	68	64	44
Wald-Chi^2-test	43***	31***	13***

Table 1: Estimation results for individual random effect regressions with dependent variable Δ_i, that is, the first difference of i's contribution in $t+1$ and t; coefficients are reported along standard errors in parenthesis; *** indicates significance at a $p < 0.01$ level, ** at a $p < 0.05$ level and * at a $p < 0.1$ level.

The estimation results show important differences between treatment conditions. The coefficients for dif_-^t and dif_+^t in TRUTH indicate an adjustment of contributions that is best summarized as conditional cooperation. That is, players contributing less than the group average tend to increase their contributions in order to (almost) match the mean contribution (see, e.g., Fischbacher & Gächter, 2010). In TRUTH, for every ECU their contribution is short of the mean, they increase their contributions by about .3 ECU. On the other hand, players contributing more than the average, tend to decrease their contributions. For every ECU their contribution exceeds the mean, they decrease their contributions by about .5 ECU.

In contrast, the second information is not provided in LIE. Rather, players in this treatment receive suspicious feedback and respond by decreasing their contributions by about 2.4 ECU. In other words, instead of adjusting the contribution (downwards) towards the mean, there is a "flat" reduction by 2.4 ECU. In turn, this means that the reduction of cooperation rates is larger under LIE than under truth if contributions are rather homogeneous (i.e., if differences are smaller than 5 ECU). However, if differences among contributions are large (i.e., larger or equal to 5 ECU), the downward adjustment of contributions in TRUTH reduces cooperation rates more severely than in LIE. Finally, providing feedback information in WARNING that is only partly ambiguous, there is no systematic reduction in contributions in this case. Thus, we cannot say much about the development in this treatment condition.

Result 2. *Contribution reductions under* LIE *are larger than under* TRUTH *if contributions are rather homogeneous, while the opposite is true for very heterogeneous contributions within a group.*

To conclude our analysis, let us assess the welfare consequences of the different feedback regimes. Hence, we analyze how players' profits develop during the game. More specifically, we analyze how profits change from the first to the second sequence of the experiment in the three treatment conditions. For this purpose, we compute the difference between profits in corresponding periods in sequence one and sequence two of the experiment. That is, the difference between mean profit in period 1 and 21, period 2 and 22 and so on shows us by which amount welfare is decreased due to shaken confidence in the feedback mechanism (LIE) or the inconvenient truth (TRUTH).

Comparing profits across treatment conditions shows that the differences between period 1 and 21, period 2 and 22 and period 3 and 23 are significantly larger under LIE than under TRUTH.[12] That is, the drop in welfare between sequence one and sequence two of the experiment is significantly higher in the initial periods under LIE than under TRUTH. Again, WARNING lies somewhat in-between the other two treatment conditions and does not show significant differences.

Result 3. *Average decrease in contributions between the initial periods of sequence one and sequence two is significantly higher under* LIE *than under* TRUTH.

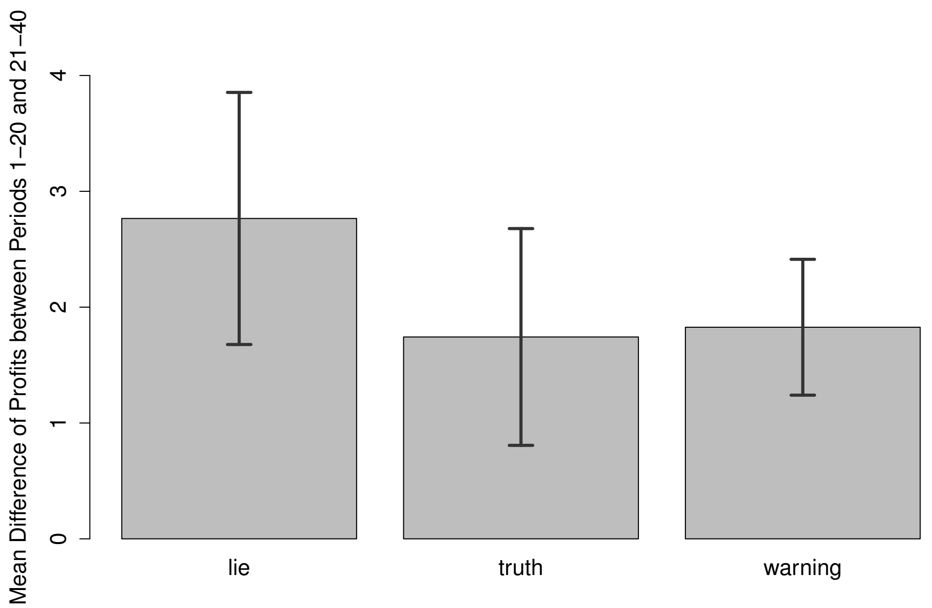

Figure 2: Mean difference in profits between first and second sequence of the experiment across treatments; error bars indicate estimations for the 95% confidence intervals.

Finally, we want to assess the overall welfare effect across both parts of the experiment. Figure 2 shows the differences between profits in the first and second sequence of the experiment across all periods by treatment conditions. The decrease varies between 1.74 ECU under TRUTH, 1.83 ECU under WARNING and 2.77 ECU under LIE. Although the treatment differences appear to be economically substantial, they are insignificantly different due to large variances within treatment conditions.[13] Therefore, we cannot provide an unambiguous assess-

[12] $p = .08$, $p = .04$, and $p = .04$, respectively, all using two-sided Wilcoxon Mann-Whitney rank sum tests.

[13] $p = .16$ for the comparison between LIE and TRUTH, $p = .23$ for the comparison between

ment here.

5 Conclusion

What goes around, comes around! Cooperation cannot be sustained either by providing reassuring lies or by telling the inconvenient truth as feedback information. That is, once the confidence in the feedback mechanism is sufficiently shaken, contribution rates decline irrespectively from the type of feedback. However, we present important qualifications for this result: there is a substantial restart effect in WARNING and TRUTH in terms of average contributions, while there is no such effect in LIE suggesting that players are more optimistic on the overall degree of cooperation in the former two treatments. Therefore, continuing with the noble lie harms initially the cooperative spirit within groups. In this sense, concerns over potential betrayals are worse than facts about betrayals.

On the other hand, receiving honest feedback harms cooperation within groups more substantially than a potential noble lie if groups are very heterogeneous in terms of their contributions. That is, there is a flat reduction by 2.4 ECU in response to a potential noble lie, whereas for every ECU the contribution exceeds the mean in TRUTH, players decrease their contributions by about .5 ECU in the consecutive period. This means that if contributions are rather heterogeneous, LIE leads to higher cooperation rates than truth. In this sense, facts about betrayals are worse than concerns over potential betrayals.

Thus, we cannot provide a clear-cut answer to the question whether it is more beneficial for society if politicians continue to tell reassuring lies or to provide the (perhaps) inconvenient truth once the noble lies have been discovered. Our third treatment does not indicate whether uncertainty concerning the size of betrayal or the actual instance of betrayal impedes the public good provision. Overall, our data favors tentatively the inconvenient truth (compare Result 3). Yet, we would like to stress that our noble lie results from a biased feedback mechanism. No human being is involved in the actual lying act (apart from the experimenter), the betrayed have no one to blame (e.g., a politician) for this act. Therefore,

LIE and WARNING, and $p = .64$ for the comparison between WARNING and TRUTH, all using two-sided Wilcoxon Mann-Whitney rank sum tests.

we have good reasons to believe that our experiment provides a rather positive estimation for the negative consequences of a noble lie. Continuing with noble lies after discovery is likely to undermine cooperation and does not enhance efficiency in democracies. In the very end, it seems that the national motto of the Czech Republic stating that “veritas vincit” – the truth prevails – in right.

References

[1] Aughey, A. (2002). The art and effect of political lying in Northern Ireland. *Irish Political Studies* 17, 1–16.

[2] Bock, O., Nicklisch, A. & Baetge, I. (2012). hroot: Hamburg registration and organization online tool. *Working Paper*.

[3] Bok, S. (2004). What's fair: Ethics for negotiators. In: Menkel-Meadow, C. & Wheeler, M. (editors), *Lies for the Public Good*. John Wiley & Sons, 371–382.

[4] Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics* 14, 47–83.

[5] Dahl, R.A. (1998). *On Democracy*. New Haven: Yale University Press.

[6] Fischbacher, U., Gächter, S. & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters* 71, 397–404.

[7] Fischbacher, U., & Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *The American Economic Review* 100, 541–556.

[8] Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10, 171–178.

[9] Hoffmann, M., Lauer, T. & Rockenbach, B. (2013). The royal lie. *Journal of Economic Behavior & Organization* 93, 305–313.

[10] Irlenbusch, B. & Rilke, R. (2013). (Public) good examples – On the role of limited feedback in voluntary contribution games. *Working Paper.*

[11] Isaac, M.R, McCue, K.F. & Plott, C.R. (1985). Public goods provision in an experimental environment. *Journal of Public Economics* 26, 51–74.

[12] Kocher, M., Cherry, T., Kroll, S., Netzer, R. & Sutter, M. (2008). Conditional cooperation on three continents. *Economics Letters* 101, 175–178.

[13] Mahon, J.E. (2008). The definition of lying and deception. In: Zalta, E.N. (editor), *The Stanford Encyclopedia of Philosophy.*

[14] Marwell, G. & Ames, R.E. (1981). Economists free ride, does anyone else?: Experiments on the provision of public goods. *Journal of Public Economics* 15, 295 - 310.

[15] Neugebauer, T., Perote, J., Schmidt, U. & Loos, M. (2009). Selfish-biased conditional cooperation: On the decline of contributions in repeated public goods experiments. *Journal of Economic Psychology* 30, 52–60.

[16] Nikiforakis, N. (2010). Feedback, punishment and cooperation in public good experiments. *Games and Economic Behavior* 68, 689–702.

[17] Pillutla, M. & Chen, X.-P. (1999). Social norms and cooperation in social dilemmas: The effects of context and feedback. *Organizational Behavior and Human Decision Processes* 78, 81–103.

[18] Walzer, M. (1973). Political action: The problem of dirty hands. *Philosophy & Public Affairs* 2, 160–180.

[19] Weimann, J. (1994). Individual behaviour in a free riding experiment. *Journal of Public Economics* 54, 185–200.

[20] Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics* 6, 299–310.

Appendix: Instructions for the TRUTH treatment

Instructions for sequence 1

Welcome to the experiment!
You will be taking part in an experiment about economic decision making. Please note that from now on and throughout the experiment **all forms of communication are strictly prohibited**. If you have a question, please indicate this by sticking your hand out of your cabin. All the decisions are taken anonymously, i.e. no other participant will know your identity. The payments at the end of the experiment will also be made anonymously, which means that no other participant will know how much your payment is.

Instructions to the experiment and general information
The experiment will consist of **2 sequences**. You will at first only receive the instructions for sequence 1. The instructions for sequence 2 will be distributed after sequence 1 has been completed. Your decisions in sequence 1 will neither have an impact on your possible choices, nor on your earnings in sequence 2.

Information about sequence 1
You are in a group consisting of 4 members in total. During all of sequence 1, you will only interact with members of your group.

Sequence 1

Contributions of group members

- In **sequence 1**, you will play 20 rounds and every round has the same structure.
- Every player receives an **endowment of 20 points** in each round.
- Every player has to decide, how many of these 20 points he or she will **contribute** to the group project.
- All points contributed to the group project will be **multiplied by 1.6** and **equally** distributed amongst all 4 players, i.e. for every point that one player contributes to the group project, **every player gets 0.4** (= 1.6/4) **points**.

- Points not contributed to the group project are kept by the individual player.

Information about the contributions of other group members

- From the second round on, you will be informed about the **average contribution** your group members made in the previous round. The average will be rounded to the next **integer number**, so that you will only see feedback in whole numbers.

Please note: Apart from the rounding, this information may <u>deviate from the actual value</u>.

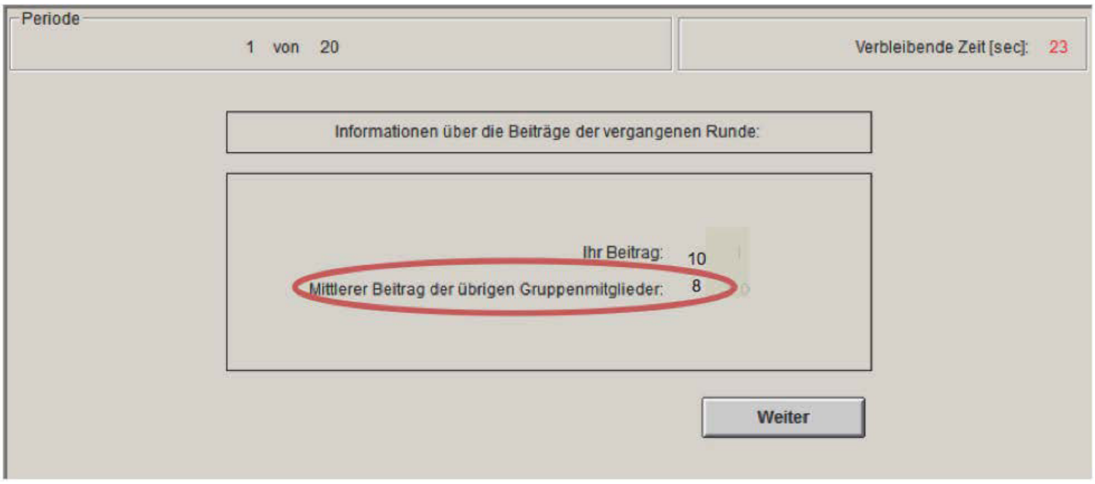

Figure 3: Screenshot - Contribution phase

Calculation of income per round

- Your income in each round will consist of two parts:
  - points, which you did not contribute
  - your ratio of the group project
- **Please note:** For the calculation of your income the actual contributions of your group members and not the announced values will be relevant.

Income per round:

(endowment (20) - your contribution to the group project)

+ sum of **all** contributions of your group $* 1.6/4$

Calculation example for the case that **you contributed 10**, while **your group members contributed 12, 8 and 4** points:

$$\begin{aligned} &= (20 - 10) + (10 + 12 + 8 + 4) * 1.6/4 \\ &= 10 + 13.6 \\ &= 23.6 \end{aligned}$$

Income in sequence 1

At the end of sequence 1, we will randomly pick **one of the twenty rounds** to calculate your income. This means that every one of the twenty rounds played is possibly relevant to your payment. Your income from sequence 1 is then the points earned in the round randomly chosen, which will be paid according to an exchange rate of 1 Euro for every 8 points.

Total income from the experiment

Your total income from the experiment will consist of the guaranteed 5 Euro, plus your income from sequence 1, plus your income from sequence 2.

Good luck!

Instructions for sequence 2

Information about sequence 2:

Sequence 2 will consist of 20 more rounds of the game played in sequence 1. The composition of your group will not change.

Please note:

In sequence 1, if you contributed **less than the average contribution** of your group in the previous round, you were shown **the actual average contribution** of your group members.

In sequence 1, if you contributed **exactly the average contribution** of your group in the previous round, you were shown **the actual average contribution** of your group members.

In sequence 1, if you contributed **more than the average contribution** of your group in the previous round, you were shown **your own contribution as**

average contribution of your group members.

However, in sequence 2 you will **always be shown the actual average contribution** of your group members.[14]

Income in sequence 2

At the end of sequence 2, we will randomly pick **one of the second twenty rounds** to calculate your income. This means that every one of the twenty rounds played is possibly relevant to your payment. Your income from sequence 2 is then the points earned in the round randomly chosen, which will be paid according to an exchange rate of 1 Euro for every 8 points.

Total income from the experiment

Your total income from the experiment will consist of the guaranteed 5 Euro, plus your income from sequence 1, plus your income from sequence 2.

Good luck!

[14] In LIE and WARNING, it is stated that "It will be the same in sequence 2," thereafter the previous paragraph is repeated replacing sequence 1 by sequence 2. In addition, in WARNING, it is stated that "However, from now on you will receive the announcement 'This information is false!' whenever, the shown average contribution differs from the actual average contribution of your group members."