Schickora, Jan Theodor:

Bringing the Four-Eyes-Principle to the Lab

Munich Discussion Paper No. 2011-3

Department of Economics
University of Munich

Volkswirtschaftliche Fakultät
Ludwig-Maximilians-Universität München

Online at https://doi.org/10.5282/ubm/epub.12160

Bringing the Four-Eyes-Principle to the Lab

Jan Theodor Schikora*†

February 17, 2011

Abstract

The 'Four-Eyes-Principle' is considered as one of the most potent measures against corruption although it lacks both theoretical and empirical justification. We show in a laboratory experiment using a standard corruption game that introducing the 4EP increases corrupt behaviour, casting doubt on its usefulness as a general recommendation. Combining data on final choices with observations on the decision making processes in teams, including a content analysis of exchanged messages, provides insights into the dynamics of team decision making and shows that the individual profit maximizing motive dominates group decision making and crowds out altruistic arguments.

Keywords: Corruption; Laboratory Experiments; Group Decision Making

JEL Classification Numbers: C72, C92, D73

*MGSE, University of Munich

†I am grateful to Martin Kocher and two anonymous referees for their helpful comments.

1 Introduction

With almost daily media attention of high profile scandals, corruption, generally defined as 'the misuse of public office for private gain' (OECD 2010), has been recognized as a major problem. In general, a corrupt transaction is illegal and exerts a large negative external effect on outsiders, which is usually assumed to be larger than the sum of benefits to the agents who are directly involved. This defines corruption as socially inefficient (Klitgaard 1988, Rose-Ackerman 1999).

In addition to traditional views of deterring an agent from engaging in criminal activity by varying the amount of penalties and the probability of detection (Becker 1968), the New Institutional Economics (NIE) of corruption concentrates on finding an institutional design that optimally exploits the instability of the corrupt transaction between a client and an official (Schulze and Frank 2003).

The instability of a single corrupt transaction stems from the enforcement problem between a bribing agent and a potentially corrupt official. Its illegal nature precludes the assistance of legal third parties, i.e. the courts (Lambsdorff 2007). The occurrence of corruption therefore relies heavily on trust and reciprocity and is difficult to explain in standard theoretical models. Nonetheless, national as well as international organizations such as Transparency International, the OECD and several national fiscal authorities publish lists of (institutional) policy recommendations containing measures to curb corruption. Along with 'staff rotation' (analyzed in Abbink 2004), the introduction of the Four-Eyes-Principle (4EP), 'a requirement that business has to be effectively conducted by at least two individuals (four eyes)', is one of the most prominent examples (Pörting and Vahlenkamp 1998, Rieger 2005, Wiehen 2005, Hussein 2005). As a result of general problematic tractability (let alone predictability) of corrupt behaviour, a theoretical analysis of the effectiveness of the 4EP does not exist. Nor is there any kind of traceable empirical evidence to support its usefulness.

Not only in the corrupt context, but also on a general level, the distinction between individuals and small groups as decision-makers has been widely ignored in the theoretical literature. Differences between the behaviour of individuals deciding alone or in a group have only recently been addressed in the field of experimental economics, where results seem ambiguous. Some studies find that the behaviour of groups is closer to standard equilibrium predictions derived from the self interested model of payoff maximization (e.g. Bornstein and Yaniv 1998, Blinder and Morgan 2005), other studies (e.g. Kocher and Sutter 2005, 2007, Cason and Mui 1997) provide experimental evidence to the contrary. Kocher and Sutter

(2007) conclude that the direction of the group decision-making effect critically depends on the nature of the task determining which of two countervailing motives, the profit maximizing motive or the competitive motive, dominates. The basic set-up of our laboratory experiment is close to those used in Abbink (2004) and Lambsdorff and Frank (2010). Our experiment is designed to assess the effects of the introduction of the 4EP on observed levels of corruption. Within this framework we model the 4EP as replacing a single official (deciding individually) with a group[1] of two officials deciding jointly according to a decision-making process that secures veto power for non-corrupt officials.

Using four different treatments we can separate two countervailing effects of the introduction of the 4EP. One is due to the difference in marginal incentives resulting from the division of the transfer between the jointly deciding officials. The other effect is determined by the group decision-making process alone (keeping marginal incentives constant). Rejecting predictions taken from (self interest based) arguments within the standard model of corruption, we find that the introduction of the 4EP increases the frequency of successful corrupt transactions unambiguously. We substantiate this hypothesis in three stages. First, we consider only outcomes (actual corruption levels). Second, we investigate behaviour in the decision-making process, i.e. we compare initial and final choices. Third, we analyse the content of electronic text messages exchanged during the decision-making process between jointly deciding officials. Our results strongly suggest the dominance of the profit maximizing motive (Kocher and Sutter 2007, Blinder and Morgan 2005). Groups reveal more functional behaviour with respect to conditional responding. By their higher (joint) cognitive capacity, groups of officials seem to be more capable of maximizing their payoffs by following strategies that are shown to lead to a higher frequency of corrupt transactions based on mutual reciprocity.

Our explanations of the observed effects are in line with the argumentation of the persuasive argument theory (Pruitt 1971). Since groups perform better in solving the enforcement problem between briber and official, the introduction of the 4EP moves behaviour further away from the theoretical prediction of selfish behaviour, which, in the corrupt context falls in line with the social optimum. Therefore our results cast doubt on the usefulness of the introduction of the 4EP and its justification as a recommended measure against corruption.

The remainder of the paper is structured as follows. Section 2 describes the experimental set-up giving details on the specifications of all four treatments. Section 3 analyses the effects of the introduction of the 4EP in the framework of the NIE of corruption and forms hypotheses. Section 4 gives details on the procedure of the experiment. In Section 5 we

[1] Although the entity consists of only two participants we call it a group rather than a team.

describe the main findings, provide a detailed explanation of the empirical strategies and interpret the results. Section 6 summarizes and concludes.

2 Experimental set-up

2.1 Corruption and the NIE

In its most general form, a corrupt transaction can be described as a Principal-Agent-Client relationship, in which the principal, represented by the government or any kind of benevolent authority, deals with its clients (private entities, e.g. firms) through its agents, the potentially corrupt (not perfectly monitored) public officials. Clients may have an incentive to transfer side-payments (i.e. bribes) to relevant officials in order to alter their behaviour with respect to their duties (i.e. fulfilling legal procedures clearly defined but not perfectly controlled by the authorities).

The main mechanism of the instable and therefore interesting part of the relationship is best explained in a simple 3-Stage game:
In Stage 1, a client (e.g. a potentially bribing firm) B decides on the level of bribe b=0,1,...,12 Experimental Monetary Units (EMU) to be given to a potentially corrupt official O. The limit of 12 EMU reflects B's budget constraint. At this point the amount of b is tripled. The factor 3 captures the idea of a difference in the marginal utilities of money between a (wealthy) client and a (poor) official (Abbink et al. 2002).

In Stage 2, O can either 'accept' the transfer b and keep the tripled amount ($3*b$) to herself or 'reject' it. In the latter case the game ends, O keeps only b to herself (as some 'benefit' from pro social behaviour) while the rest (the remaining amount of $2*b$) is used for the 'public benefit'.[2] If O accepts the bribe, she gets the tripled transfer for sure. She automatically enters Stage 3 where she decides between two options. The first option includes initiating an increase of B's payoff by 16 EMU (delivering the corrupt service) at the fixed costs of 4 EMU (to herself). In this case, the 'public' suffers substantially (by −24 EMU). Note that independent of the size of the transfer, the negative externality is always larger than the sum of payoffs for the agents B and O so that a successful corrupt transaction is always inefficient by construction.

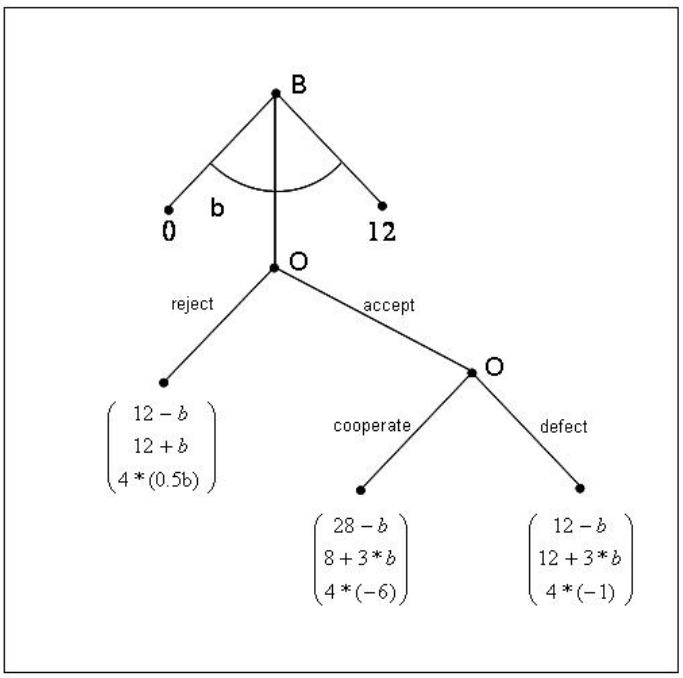
In the second option O arranges nothing (implicitly defects on B), saving costs for herself

[2] This implies the assumption of equal marginal utility between O and the public.

and the public from the negative externality. The costs of delivery may be interpreted as the certainty equivalent of the lottery: punishment in the case of detection and corrupt success, or as the practical costs of engaging in a criminal activity, e.g. hiding illegal activities from colleagues and superiors.[3]

If O decides against delivering the corrupt service in Stage 3, B does not get the bonus, O does not bear the costs of 4 EMU, and only a minor negative externality to the public (4 EMU) is realized. Within the standard self interested model it is never optimal, neither in a one shot nor in a finitely repeated game, for O to deliver the corrupt service and hence for B to transfer a positive bribe, see Appendix A for a short proof. Figure 1 represents the basic set-up in its extensive form.

Figure 1: Extensive form representation



2.2 The 4EP

In order to investigate the effects of the introduction of the 4EP on the level of corruption in this set-up, we consider two participants instead of one making the decisions of the official

[3]Abbink (2004) uses a fourth stage which accounts for probability of detection. Using a lottery instead of a fixed amount to model the cost of corruption adds another problem of individual differences (risk aversion). This would require to disentangle potential treatment effects from differences in risk aversion.

in Stage 2 and Stage 3. We define the level of corruption as the frequency of successful corrupt transactions relative to the total number of possible transactions.

We use four different versions of the game as treatments of the experiment. In all four treatments, subjects play their version of the corruption game for ten successive periods. After each period they learn about their group partners'/partner's (payoff relevant) choices.

All treatments are run in a partner design so that all subjects remain in their respective unit (of B and O subjects) for all ten periods of the experiment.

In our experiment, the 'public' is modelled in two different ways. In one set of sessions (mode 1) we model the externality on the public as payments (reductions of payments) to four randomly chosen participants of the (same session of the) experiment. In the second set of sessions (mode 2) we model the externalities as increases or decreases in the amount of a donation to the public aid organization 'Doctors without Borders'. The total amount of added or deducted payments is equal across the two modes.

IDT1[4]

In Treatment 1, the 'Individual Decision-making Treatment 1' (IDT1), we consider units of two subjects, one in the role of the official O and one in the role of the potential briber B. The 3-Stage corruption game (see Figure 1) is played for ten consecutive periods. At the end of each period, all participants get to know their own payoffs. Additionally, type B subjects get information about the (Stage 2 and Stage 3) decisions of their transaction partners (of type O). While through this information all subjects know about the negative or positive externalities they have helped to cause to the public (four randomly chosen participants in mode 1 or 'Doctors without Borders' in mode 2) they do not learn about the magnitude of the spill-overs that may have been caused to them by the decisions of the subjects outside their unit in mode 1.

TDT1

In the second treatment, the 'Team Decision-making Treatment 1' (TDT1), we form units of three subjects, one B and two O types. The B type decides in Stage 1 about her bribe b which is tripled and then transferred to both officials of her (3-player-)unit. Note that although the amount goes to two players, it is subtracted only once from B's account. The parameters of the game are set in such a way that the incentives for the officials are equal to the ones in IDT1, given the amount of bribe. This way we can separate the true 'Group Decision-making Effect' (GDE) from effects stemming from the partition of the bribe between the two officials.

[4]See Figure 6 in Appendix B for the extensive form representation.

In Stage 2 and Stage 3, the two officials of a unit make their decisions jointly. In both stages they decide independently first. If they do not come to an unambiguous decision (e.g. one official decides for 'reject', the other for 'accept' in Stage 2), they learn about each other's choice and decide again. If there is still no agreement, they get the opportunity to communicate with each other via a real time electronic 'chat' in which they can, for one minute, exchange electronic messages (see the translated instructions in Appendix C).

If there is still no mutual consent, the corruption-unfriendly choice is taken ('reject' in Stage 2 and 'defect' in Stage 3). This rule reflects the veto-power of officials who do not want to support a corrupt transaction.

IDT2

The 'Individual Decision-making Treatment 2' (IDT2) differs from IDT1 only in the number of possible transactions between a particular B-O pair. In this treatment we consider units of four, two type B and two type O participants. Every type B participant sends only one transfer to one of the two officials in her unit per period. This means that playing the game for ten periods makes five possible transactions per pair, producing two transactions per period and four-player-unit.

TDT2

In the 'Team Decision-making Treatment 2' (TDT2) we form again four-player-units. Each of the two type B players sends one transfer each to the group of two officials who decide jointly in Stage 2 and Stage 3, according to the same decision-making process explained for TDT1. Contrary to the case of TDT1, the transfer is split equally so that each of the officials receives only half of the tripled amount of the transfer chosen by the respective briber ($3 * 0.5 * b$). This means that each of the type B participants makes one decision, while each of the two officials has to decide in two separate situations per period. As a consequence, each type O subject receives two payoffs per period, while B receives only one.

Here the parameters are set in such a way that one transaction in TDT2 corresponds to two transactions in IDT2 in terms of payoffs for type O subjects. This means that O's individual incentives in a certain situation (transaction) are not equal to those in IDT1, TDT1 or IDT2, since a certain transfer b leads to double the amount of revenue reaching a single official in IDT2 as compared to the revenue reaching each of the subjects in the role of the official in TDT2. Not only gains (tripled transfers) but also costs (4 EMU) are shared equally between the two officials.

2.3 Related literature

The most important difference between the gift exchange game (Fehr et al. 1993) as well as the investment (trust) game (Berg et al. 1995) and the corruption game is that cooperation in the form of positive reciprocity is beneficial for all members of society (i.e. the society as a whole), whereas in the case of corruption, cooperation is only beneficial for the client and the official but not for the public in general (and those who are hit by the negative externality in particular). The negative externality is assumed to be high enough to result in a net social loss. This difference may have strong effects on the level of cooperation and positive reciprocity, as well as on its behavioural explanations and motivations.

Contrary to theoretical predictions applying the standard self interested model of payoff maximization, Abbink et al. (2002) find a large amount of cooperation in a series of laboratory experiments using one-shot as well as (finitely) repeated versions of set-ups comparable to ours.[5]

These findings can neither be fully explained (or predicted) by the standard self interested model (see Appendix A), nor by models of reciprocity based on social preferences (see e.g. Fehr and Schmidt 1999, Charness and Rabin 2002). But even if there was a theoretical explanation of observed corrupt behaviour there would be a lack of predictive power for the effect of group decision-making.

We know from the experimental literature on group decision-making that decisions made in a small group or team can differ substantially from individual decisions even if individual marginal incentives are equal (Chalos and Pickard 1985, Levine and Moreland 1998). Experimental evidence suggests that the direction of the effect of group decision-making is ambiguous and depends highly on the nature of the particular situation. The majority of studies find that decisions made in small unitary groups (which is the case in our study) act more in line with the predictions of the self interested model of payoff maximization (Blinder and Morgan 2005, Bone et al. 1999, Bornstein and Yaniv 1998, Kugler et al. 2007), while there is also contrary evidence (e.g. Cason and Mui 1997). Kocher and Sutter (2007) show (using a gift exchange game) that decisions made by a small group may be either more or less in line with selfish preferences.

[5]The common feature is the trade-off between reciprocity based maximization of individual long-term payoffs and a combination of impulses of myopic payoff-maximization and preferences for social efficiency.

3 Hypotheses

In our analysis, we distinguish between two main effects of the introduction of the 4EP with respect to the officials' behaviour. First, the introduction of the 4EP will share a bribe between two officials instead of going to just one. Keeping B's behaviour (i.e. the amount of transfer) constant, the splitting of the bribe causes each official to receive half of what a single official would have got. We call the officials' immediate reaction to the lower benefit from a bribe in TDT1 and TDT2 the 'Bribe Splitting Effect' (BSE). Second, we consider the pure effect of group decision-making when we hold marginal effects constant and call it the 'Group Decision-making Effect' (GDE).

3.1 Bribe Splitting Effect

From a series of experiments using comparable set-ups (e.g. Abbink 2004), we know that the probability of success increases with the level of transfer (bribe). For simplicity we assume that the bribe is shared equally between the two officials, ignoring potential distributional issues. The existence and magnitude of an effect stemming from the splitting of a bribe depends on what officials condition their behaviour on. Officials may consider the monetary benefits of the bribe only or they may include the 'intentions' of the briber and their own moral responsibility.

If subjects condition their behaviour on intentions and equilibrium outcomes alone (i.e. they consider the 'kindness' of B's decision only in the sense that it leads to a certain outcome, given that the transaction is successful) we would not predict any difference between the conditional behaviour of type O subjects in TDT1 and TDT2. Due to the construction of the experiment[6], strategies leading to equalized outcomes between B and O (ignoring the negative externalities to the public) are the same across treatments and therefore require the same actions for both types. Hence outcome-based models of inequity aversion (e.g. Fehr and Schmidt 1999) would yield the same predictions across treatments assuming the irrelevance of the negative externality.

Hypothesis 1: "Holding bribe levels constant, there will be no difference in corruption levels between TDT1 and TDT2, if officials condition their reciprocal behaviour exclusively on intentions or consider the equalization of payoffs only."

[6]Including the difference in the number of transactions played per period by the different types.

Hypothesis 1 will be rejected if the actual amount of bribe in a particular situation has an effect on the probability of success of a corrupt transaction (being different in TDT1 as compared to TDT2). In this case we call the effect the Bribe Splitting Effect (BSE).

3.2 Group Decision-making Effect

In order to measure the effect of group decision-making (GDE) separated from BSE, we have to compare the behaviour of subjects deciding alone and subjects deciding within a group in situations in which all relevant decision-makers face the same marginal monetary incentives. The comparison between IDT1 and TDT1 satisfies this condition. So we compare the expected number of successful corrupt transactions, $E(N_{success})_{IDT1}$ (one official decides individually) with the expected number of successful corrupt deals, $E(N_{success})_{IDT1}$ (two officials decide jointly) facing the same bribe b.

In IDT1, the probability of success of a corrupt transaction is $p_i(b)$ for the deal in which official O_i is relevant. Assume that the probability of corrupt success (reciprocity) is positively dependent on the relevant bribe b (see e.g. Abbink 2004).
In TDT1, officials O_i and O_j decide jointly in Stage 2 and Stage 3. The group decision-making process provides veto power for non-corrupt and non-reciprocal behaviour ('reject' in Stage 2 and 'defect' in Stage 3).
In this case the probability of success is $p_i(b)p_j(b)$ (both officials have to decide in favour of corruption), if decisions are completely independent. Since $p_i(b_1) \leq 1$ and $p_j(b_2) \leq 1$, $E(N_{success})_{IDT1} = p_i(b) \geq p_i(b)p_j(b) = E(N_{success})_{TDT1}$. As long as the individual behaviour of officials is independent of the decision-making process (including the observation of or the belief on the behaviour of the other official), the expected number of successful corrupt transactions should be weakly greater under IDT1 than under TDT1.

Hypothesis 2: "If decisions are completely independent of the decision-making process we will observe lower (relative) numbers of successful corrupt transactions in TDT1 than in IDT1."

Experimental evidence shows that individual decisions are far from independent when made inside a group or team. The decisions made by small groups tend to be more in line with the predictions of the self interested model of payoff maximization when considering bargaining situations in which competition plays a relevant role, while in games representing social dilemmas, e.g. the public goods or the gift exchange game, groups may even move further away from the predictions of standard game theory (Cason and Mui 1997, Levine and Moreland 1998). Kocher and Sutter (2007) track this back to two opposing motivations driving the effect of group decision making, the competitive and the profit maximizing motive. The total effect depends on which of these dominates.

The set-up of a corrupt transaction cannot easily be categorized into one of the two situations. Where reciprocity-based cooperation increases the level of social welfare in most applications (as it usually helps to overcome a social dilemma), corrupt reciprocity decreases social welfare (and efficiency) by design. For the official it is always individually (in the short term) as well as socially optimal not to reciprocate in the corruption game.

Competitive Motive

The pure fact of group membership may cause a shift in individual preferences towards a decision that reflects higher awareness of competition with other groups or individuals (minimal group paradigm, Tajfel and Turner 1986). In our situation this may push groups to behave more in line with the predictions of standard game theory. Members of a group may follow strategies that increase the difference of pay-off levels between in- and out-group members. Especially under mode 1 (fellow subjects are hit by the negative external effect), this may help cooperation within and hinder cooperation across groups[7].

What distinguishes the individual from the group decision-making treatments in our experiment is the additional in-group-effect between jointly deciding officials in TDT1 and TDT2. The creation of a sub-group by letting officials decide jointly may result in more competitive behaviour towards their type-B transaction (unit) partners, which may result in myopic profit maximization at the cost of corrupt reciprocity. We denote this motivation as the Competitive Motive (CM).

Profit Maximizing Motive

The *Profit Maximizing Motive* (PMM) causes groups to make decisions that yield larger payoffs (in the long run) while, if necessary, shifting behaviour even further away from equilibria predicted by standard theory (Kocher and Sutter 2007). Despite being inefficient on a social level, a successful corrupt transaction yields the largest individual payoffs for the transaction partners (unit), given the behaviour of other groups. Groups may be more capable of suppressing short-sighted impulses of behaviour which may maximize myopic payoffs but ultimately decreases total individual payoffs of all transaction partners. This behaviour

[7] This may involve the understanding of participants being involved in a reverse public goods dilemma.

includes free-riding or defecting in social dilemmas (e.g. the public goods game using the voluntary contribution mechanism) and failing to foresee the breakdown of future cooperation (reciprocal relationships). The Persuasive Argument Theory (PAT, see Pruitt 1971, Bishop and Myers 1974, Burnstein et al. 1973) predicts that groups are more successful in finding strategies that maximize their members' long term payoffs. Explorative and knowledge capacities in groups are expected to be greater than those of a single individual. Chalos and Pickard (1985) proclaim that groups are better in processing information load. In games where payoff maximizing strategies are as complicated as in the repeated corruption game, we expect groups to develop and follow more successful strategies than individuals with respect to maximizing their members' monetary payoffs when we assume that groups and individuals exhibit equal preferences with respect to the trade-off between individual and social welfare maximization.

Hypothesis 3: "If the group decision-making process is dominated by the CM, outcomes in TDT1 will be closer to the game theoretical predictions than those in IDT1. If the PMM dominates group decision-making, groups will produce higher levels of corruption by following strategies that are more successful in maximizing their members' individual payoffs."

3.3 *B*'s behaviour

The introduction of the 4EP may not only affect the behaviour of the officials but also that of the bribers. The direction of the effect depends entirely on the beliefs about the (effects on the) behaviour of the official(s). Relying on the assumptions of standard game theory (within the self interested model), we do not expect bribers across treatments to adhere to different beliefs about the behaviour of groups and individuals, since in all treatments the Sub game-perfect Nash Equilibrium (see Appendix A) is unique and predicts neither positive transfer levels nor positive reciprocation.

While predictions on B's expected reaction on the anticipation of the GDE are difficult, we can form hypotheses on the direction of the effect stemming from B's anticipation of the BSE. This can be quantified by comparing average bribe levels between TDT1 and TDT2. If bribers anticipate the BSE (correctly) they may send larger transfers in order to compensate the splitting of the bribe.[8]

Hypothesis 4: "The bribe level (and distribution) will be different in TDT1 and TDT2 if type B subjects anticipate officials to behave according to the BSE and react accordingly."

[8]Note that such a reaction can not be explained by models of inequity aversion, since a strategy that aims at equalizing payoffs would not proclaim different levels of transfers across these treatments.

3.4 Total effect

In order to evaluate the usefulness of the implementation of the 4EP, we consider the total effect, i.e. the combination of the BSE and the GDE. The total effect can be directly measured by the comparison of IDT1 (or IDT2) and TDT2. The introduction of the 4EP can only help to reduce the level of corrupt activity, if the conditions of either of the two following situations hold. First, the 4EP will certainly reduce corruption, if the BSE and the GDE are both positive. Second, it will reduce corruption, if a positive BSE over-compensates a negative GDE (which is dominated by the Profit Maximizing Motive). The introduction of the 4EP is counter-productive even without considering the costs of the installation of such an institution if the GDE is negative and over-compensates the BSE.

Hypothesis 5: "If the GDE is negative (PMM is stronger than CM) and dominates the BSE, the average rate of corrupt success will be greater in TDT2 than in either IDT1 or IDT2. In this case the 4EP is counter-productive."

4 Procedure

All 8 sessions (two sessions for each treatment) were programmed and conducted at the experimental laboratory MELESSA at the University of Munich. It used the program Z-Tree (Fischbacher 2007) and the organizational software Orsee (Greiner 2004). Each session was conducted with 24 subjects (a total of 192 participants). Subjects were randomly assigned a type, (B or O) and randomly allocated into units of two in IDT1, into units of three (one type B and two type O participants) in TDT1 and into units of four (two type B and two type O participants each) in IDT2 and TDT2. In all treatments group members stayed together in their units for all 10 periods (partner design) where full anonymity was ensured.[9] Every period was paid where 1 EMU was worth 5 eurocents. Payoffs were summed up over all 10 periods and paid out in private at the end of the experiment. The whole experiment took less than 90 minutes. The instructions were kept completely neutral, avoiding any language indicating the subject of research in order to concentrate on the specific features of the model and minimize the differences between the instructions of the treatments. Abbink and Hennig-Schmidt (2006) show that framing has no significant effect on behaviour in the corruption game. All this was common knowledge to all participants. Understanding of the (rather complicated) set-up was insured by partly reading out the

[9]Note that the interaction of officials was conducted via a chat which did not allow for any form of identification, see the Appendix C for the full instructions of TDT2.

instructions, answering questions in private and checking of several control questions. At the end of the experiment subjects filled in a questionnaire including demographic information. Payoffs lay between 4 Euros and 25 Euros excluding a show up fee of 4 Euros which is standard to experiments at MELESSA. Average earnings amounted to 14.37 Euros. A total of 249.10 Euros was paid out as a donation to the organization 'Doctors without Borders' as a result of the decisions made by the participants in the treatments where we chose mode 2 as a model of the negative externality (one session in IDT2 and one session in TDT2).

5 Results and Interpretation

The negative external effect has been modelled in two different ways. According to predictions using the standard self interested model, there should be no effect of either of the two models and hence no difference between the set-ups. The interaction via (unknown externalities) between groups may however produce an unwanted additional in-group effect and destroy the idea of the external effect as an unreciprocated reduction in the payoff of unrelated third parties. Corrupt behaviour might be considered as a payoff equalizing equilibrium ('super-game'). In order to rule out effects stemming from these considerations we applied mode 2 in one session of IDT2 and one session of TDT2. In mode 2, the 'super-game' problem is eliminated by modelling the negative externality as a reduction of a (fixed) amount of donation to the public aid organization 'Doctors without Borders'.

Applying (pair-wise) two-sided Mann Whitney U-tests[10] we compare all relevant variables, i.e. the average total transfer-level, the transfer levels after success and failure of a corrupt deal (measure of the client's reciprocity), the average relative number of successful deals, the percentage of rejected bribes and the percentage of zero-value transfers between observations of the two modes and find no significant difference ($p \geq 0.363$; $N \geq 16$). We therefore pool the data from these sessions in the respective treatments for the entire analysis. The absence of a 'super-game' effect may be explained by subjects having difficulties in forming beliefs of higher orders (see Anderson and Holt 1997, Hung and Plott 2001).

[10] Unless stated otherwise (exact, highest or lowest) $p-$values and numbers of observations (N) apply to the two-sided Mann Whitney U-test.

5.1 Descriptive results

Table 1: **Performance variables**

	Corrupt success		*Payoff*		*Neg. Externality*		*Transfer level*	
Treatment	mean	std.dev.	mean	std.dev.	mean	std.dev.	mean	std.dev.
IDT1	0.25	0.09	13.13	3.92	2.90	1.01	3.18	1.52
IDT2	0.25	0.06	12.72	3.78	3.49	2.12	3.16	1.25
TDT1	0.47	0.07	7.32	4.26	5.13	2.91	3.58	1.15
TDT2	0.42	0.05	9.40	4.05	5.38	1.97	4.15	1.24

All means are calculated as averages across periods and (relevant) participants of the respective treatment.

Table 1 shows the values of the four main performance variables in all treatments. *Corrupt success* depicts the average share of successful transactions per unit ($\frac{N_{success}}{N_{total}}$).[11] *Payoff* represents the average payoff level (in Euros) after the reduction of the negative externality[12] (final payoff) per subject. *Neg. Externality* describes the level of the (relevant) negative externality in Euros. *Transfer level* measures the average amount of bribe (in EMU) transferred by type B participants per period and unit.

5.1.1 Corruption levels, payoff and externalities

Comparing corrupt success rates (*Corrupt success*) between treatments, we can identify a relatively small but significant Bribe Splitting Effect. The difference in corrupt success rates between TDT1 and TDT2 (0.05) is significant ($p = 0.082$; $N = 28$). We reject **Hypothesis 1**. Corrupt success does not seem to depend exclusively on intentions or final outcomes. Moreover we find a substantial Group Decision-making Effect (**Hypothesis 2**). The negative difference in the corrupt success-levels between IDT1 and TDT1 amounts to 0.22 ($p = 0.034$; $N = 40$) and strongly suggests the dominance of the Profit Maximizing Motive (**Hypothesis 3**). The large and significant difference (0.17) between IDT1/IDT2 and TDT2 (IDT1 vs. TDT2: $p = 0.002$; N=36, IDT2 vs. TDT2: $p = 0.042$; N=24) indicates a negative total effect of the introduction of the 4EP even if we control for the difference in

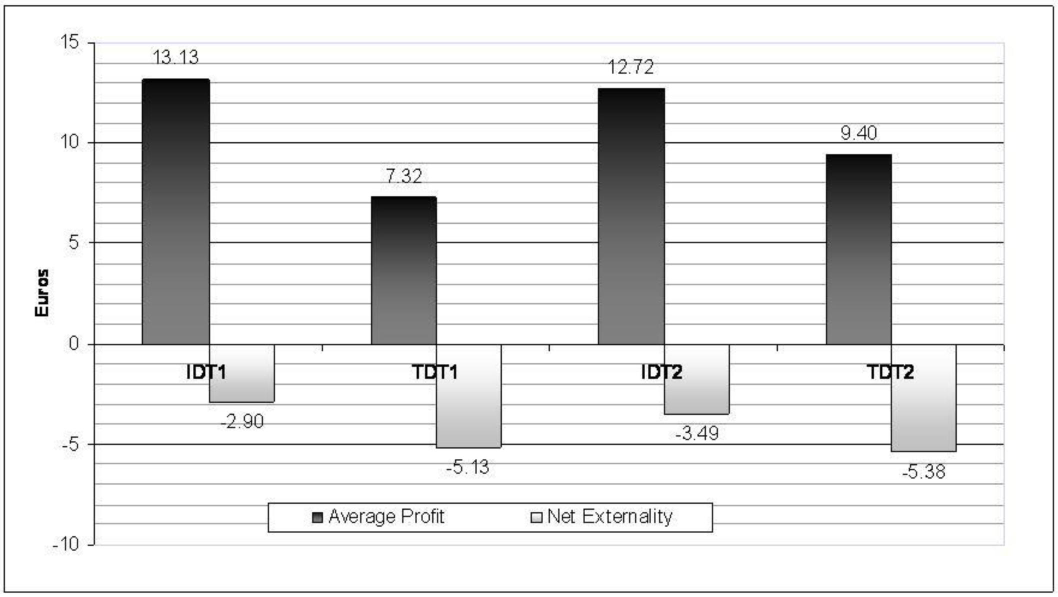
[11] Note that units contain two, three or four individuals, which makes a comparison between absolute levels of corrupt success inconclusive.

[12] To be able to accurately compare payoff levels between all treatments, we subtract the relevant share of the reduction of the donation from the actual payoff in the externality mode 2.

the number of repetitions (between IDT1 and TDT2) confirming **Hypothesis 5**[13].

The average net payoff[14] is more meaningful in terms of a welfare comparison between the treatments than the level of corruption. Figure 3 shows the average total payoff level (*Payoff*) per subject in comparison to the average negative external effects (caused by the average subject, *Externality*) for all four treatments.

Figure 2: Payoff levels and levels of externalities



The differences in the individual profitability from engaging in corrupt activities between the treatments are not as large for the officials. We estimate a simple OLS regression, measuring the marginal effect of N, the number of successful corrupt transactions (number of choices in which the (group of) official(s) has cooperated) on 'Payoff', the total payoff of each (group of) official(s).[15]

$$M2: \quad \text{Payoff}_i = \beta_0 + \beta_1 N_i + \gamma D_i + \delta D_i * N_i + \epsilon_i \qquad (1)$$

Vectors D and $D * N$ have analoguous interpretations as D and $D * b$ in Model M1. The results of the regression are reported in Table 9 of Appendix 2B. Using OLS as well as Tobit (as a robustness check), we find a strong positive effect of the number of successful transactions on the total payoff in all treatments. The profitability of being corrupt is

[13]The Profit Maximizing effect dominates the Competetive Motive in group decision-making. The resulting effect is stronger than the Bribe Splitting Effect.

[14]As net payoff we define the sum of payoffs after subtraction of the externality caused by others' decisions.

[15]Since officials within a unit in the GD treatments decide jointly they receive the same payoff and are therefore treated as a single observation. Officials who are in the same unit but decide independently (IDT2) are treated as individual observations but we cluster their standard errors in the regression.

significantly higher in both GD treatments than in the ID treatments (t-tests for IDT1 vs. TDT1/TDT2 and F-tests for IDT2 vs. TDT1/TDT2; $p \leq 0.004$). While officials in the ID treatments earn on average only 15.892 (IDT1) and 15,075 (IDT2) EMU more for an additional successful corrupt transaction, the rate is at 19,035 (TDT1) and 20.923 (TDT2) EMU considerably higher in the GD treatments (Differences between ID and TD treatments are all significant; t-test for IDT1 vs. TDT1/TDT2, F-test for IDT2 vs. TDT1/TDT2: $p \leq 0.001$).

5.1.2 Transfer levels

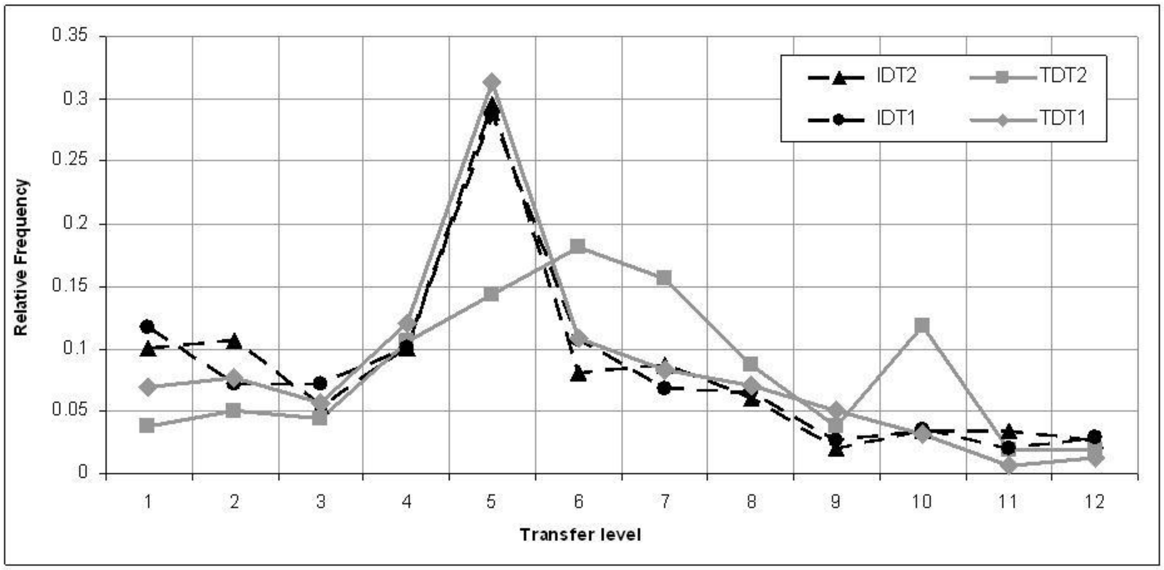
Average bribe levels (including 0-transfers) are substantially (and significantly, U-tests in all treatments: $p < 0.001$, $N \geq 12$) larger than 0 for all treatments and almost identical within the ID treatments (3.18 and 3.16 EMU). Transfers are at 3.58 only insignificantly larger in TDT1 than in the individual decision-making treatments (TDT1 vs. IDT1/IDT2: $p = 0.351/0.464$, $N = 36/28$). At 4.15 EMU, the average transfer level in TDT2 is significantly (U-tests: TDT2 vs. TDT1/IDT1/IDT2; $p \leq 0.041$; $N \geq 24$) larger than those in any of the other treatments. The large difference in transfer levels between TDT1 and TDT2 suggests that bribers anticipate different behaviour from officials and react accordingly. Taking into account that success levels in corruption are significantly lower in TDT2 than in TDT1 despite the positive difference in transfer levels, we conclude, assuming realistic beliefs, that bribers anticipate the BSE and 'react' by trying to 'convince' officials by transferring larger bribes (**Hypothesis 4**).

The distribution of the size of transfers reveals even more information about B's behaviour. Figure 3 shows the relative frequency of transfer levels for all treatments. Transfers are almost identically distributed in IDT1, TDT1 and IDT2. There are only few low ($b < 4$ EMU) and high ($b > 8$ EMU) transfers. We observe a very strong mode at $b = 5$. This particular observation may, e.g. be explained by subjects behaving according to preferences based on inequity aversion (Fehr and Schmidt 1999). The strategy [$b = 5$ EMU; 'accept'; 'cooperate'] leads to equal payoffs for B and O within a unit in all four treatments.[16]

The distribution in TDT2 depicts a significantly different pattern. We compare the distribution of bribes in TDT2 to those in all three other treatments with a Kolmogorov-

[16]For the results to be explained by social preferences we either need to assume that B and O's reference group excludes the public (other participants in mode 1 or recipients of donations from 'Doctors without Borders' in mode 2), or assume a certain structure of beliefs on the (corrupt) behaviour of the other units (only valid for mode 1).

Figure 3: Relative distribution of transfer levels

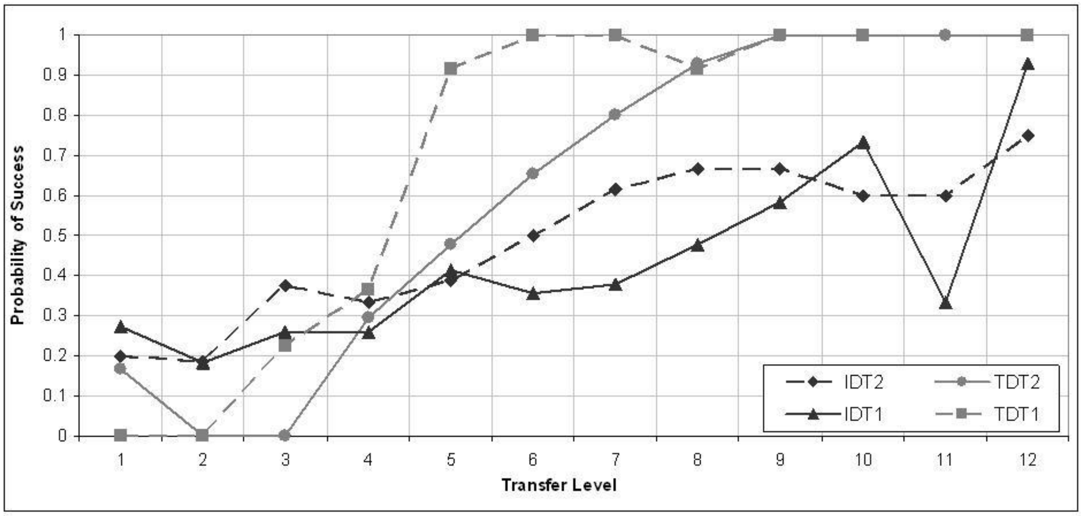


Smirnov test ($p < 0.001$; observations of strictly positive bribes: $N \geq 382$; all observations: $N = 480$). In TDT2 probability mass is shifted towards the higher end of transfer levels. The second mode at $b = 10$ EMU can be explained by a reaction to the bribe splitting. Within a certain situation between a particular briber and two officials, the strategy [‘$b = 10$ EMU’; ‘accept’; ‘cooperate’] leads to equalized payoffs. However, since the total payoff for a certain period consists of two payments for each of the officials, while the briber only receives one, this strategy does not equalize outcomes with respect to the total period payoff. It is common knowledge (and was made explicitly clear with the help of several control questions in the instructions) that it is the strategy [‘$b = 5$ EMU’; ‘accept’; ‘cooperate’] that yields equal outcomes (in expectations) for all transaction partners, just as in the other treatments. While the monetary benefits as well as the monetary costs are split in TDT2 (compared to TDT1) we may interpret the higher levels of transfer levels in TDT2 as a premium compensation for the ‘moral’ costs of causing damage to the public, which applies to both participants, since they both have full (moral) responsibility for the corrupt outcome.

5.2 Conditional reciprocity

Throughout the experimental literature on trust and reciprocity in general (Fehr et al. 1993) and on the corruption game in particular (see e.g. Abbink et al. 2002, Abbink 2004, Lambsdorff and Frank 2010), the scale of reciprocation has been found to depend critically on the first mover's behaviour (in our case the level of b). In order to explain differences in the level of corruption between our treatments we need to control for the level of transfer. Figure 4 shows the probability of a successful corrupt deal (in %) for any positive bribe level (1-12 EMU) for all treatments.

Figure 4: Success rates conditional on Transfer



In all treatments success rates increase with the level of b (Spearman rank correlation coefficients $\rho \geq 0.85$; $p < 0.001$ for all treatments). This is not surprising, since the cost of corruption is fixed, and future gains from successful reciprocity increase with the size of the bribe, assuming that bribe levels are positively correlated across periods. While the relationship between success and transfer levels seems almost linear in IDT1 and IDT2, we observe a different pattern for the GD treatments. In these treatments the conditional probability is substantially higher for large transfers ($b > 5$ EMU) and slightly lower for small ones ($b < 5$ EMU). We find significant positive differences in probabilities between GD and ID treatments for larger transfer levels $5 < b \leq 11$, considering transfer levels separately (TDT1 vs. IDT1/IDT2: $p \leq 0.021$, TDT2 vs. IDT1/IDT2: $p \leq 0.094$; $N \geq 12$). The negative differences for $b < 5$ are not significant for pair-wise comparisons except for $b = 2$ (IDT1/IDT2 vs. TDT1/TDT2: $p \leq 0.08$; $N = 22$). Differences at $b = \{0; 1; 3; 4; 5; 12\}$ are not significant at any relevant level between ID and TD treatments ($p \geq 0.127$; $N \geq 5$).

Parametrically, the most straight-forward way to quantify the differences in the probability of a successful corrupt transaction conditional on the relevant transfer between the treatments is to use a linear panel regression (random effects) controlling for clustered standard errors on the unit level. Since we are primarily interested in the causal relationship between the level of transfer (b) and the success levels (SC), we do not distinguish between

a corrupt deal that failed in Stage 2 or in Stage 3. Treating the decisions 'reject' and 'defect' equally with respect to the outcome of a corrupt deal (success or failure), we do not have to take the selection process of reaching Stage 3 into account.[17]
We use the following specifications for the linear probability model:

$$M3: \quad \text{Prob}(SC_{it} = 1|\psi X) = \beta_0 + \beta_1 b_{it} + \gamma D_i + \delta D_i * b_{it} + \theta Z_i + \epsilon_{it} \tag{2}$$

ψ stands for the vector of coefficients. X represents independent variables. Again, vectors D and $D * b$ stand for treatment dummies and interaction terms of treatment dummies with the transfer b, just as in Model M1. Vector Z contains individual demographic characteristics (e.g. age, gender[18], an interaction term between gender and the level of transfer, etc.) obtained from the questionnaire. Since we do not find any significant effects with any of these characteristics we do not report them in the regression output (Table 3).[19] M3 in Table 2 reports the results (coefficients and standard errors) of the linear probability model.

In all treatments, we find that an additional unit in transfer (b) increases the probability of the corrupt success significantly (1%-level). The effect is significantly stronger in both GD than in the ID treatments (t-tests for IDT1 vs. TDT1/TDT2: $p \leq 0.003$, F-tests for IDT2 vs. TDT1/TDT2: $p < 0.001$).

[17] Treating the outcomes of 'reject' and 'defect' differently would require a Heckman-selection process explaining the selection of cases in which Stage 3 is reached (Heckman 1979).

[18] For officials in the GD treatments we use a dummy for 'all-female' groups and do not distinguish between 'all-male' and mixed groups.

[19] As a robustness check we ran the panel regression with a series of specifications, including a regression excluding Z and a set of pooled OLS regressions including dummy variables for periods. None of these specifications yield results qualitatively different from those reported in the left part of Table 2.

Table 2: **Output of (M3) and (M4)**

	Dependent variable: SC			
	(M3) Lin. Prob		(M4) Probit	
	Coefficient	Stand. error	Coefficient	Stand. error
Constant	0.0582^{***}	0.0215	1.5832^{***}	0.1052
D_{IDT2}	0.0981^{***}	0.0351	0.5442^{***}	0.1536
D_{TDT1}	-0.0343^{**}	0.016	-0.4637^{**}	0.2091
D_{TDT2}	-0.0722^{***}	0.0232	-1.3272^{***}	0.3162
b	0.0593^{***}	0.0132	0.2123^{***}	0.0171
$D_{IDT2}{}^{*}b$	-0.0208	0.0175	-0.0981	0.0872
$D_{TDT1}{}^{*}b$	0.0521^{***}	0.0162	0.2438^{***}	0.0402
$D_{TDT2}{}^{*}b$	0.0462^{***}	0.0133	0.3402^{***}	0.0540
	Pseudo $R^2 = 0.36$		–	

*** denotes significance at the 1%-level.

Number of subjects: 96, Number of clusters: 64, Number of periods: 10

The non-linear relationship of success-probabilities and transfer levels observed in Figure 4 can be quantified by a simple maximum likelihood model. To account for differences in the marginal effect of an additional unit in transfer on the success probability across transfer levels we run the following Probit model in its panel version (random effects).[20] We use the same set of independent variables and repeat all robustness checks (pooled version etc.) applied to the linear probability model (M3).

$$M4: \quad \text{Prob}(SC_{it} = 1|\psi X) = \phi(\beta_0 + \beta_1 b_{it} + \gamma D_i + \delta D_i * b_{it} + \theta Z_i) \tag{3}$$

Again, ψ stands for the vector of coefficients and X for independent variables. As expected, qualitative results (direction and significance of the evaluated marginal effects at the mean of of transfers $\bar{b} = 3.46$) do not change compared to the results from the linear probability model, see (M4) in Table 2. Table 8 in Appendix B reports marginal effects of the relevant[21] variables as well as predicted conditional probabilities of success of model M4. The Probit model shows that marginal effects are lower in the TD than in the ID treatments for low

[20]See Pereira et al. (2006) and Gneezy and List (2006) for examples of the use of a panel version of maximum likelihood models in comparable settings, i.e. repeated gift exchange games.

[21]Again we do not report any coefficients that are not significant, e.g. a dummy variable for gender.

transfers, $b < 3$, while they are higher for $b \geq 4$. Consequently, the predicted success levels (probabilities) conditional on the transfer level are lower in the TD treatments than in the ID treatments for $b \leq 4$ while they are larger for $b \geq 6$ (see Table 8 in Appendix B).

The pattern shown in Figure 6 and quantified in M4, i.e. a stronger curvature of the probabilistic cumulative distribution function for GD than for ID treatments, may be explained by differences in the strategies between groups and individuals. On the one hand, groups of officials seem to 'defect' (or 'reject') more often in the case of low transfer levels. On the other hand, they seem to be more likely to reward high transfers than their individual counterparts by corrupt reciprocity. We interpret this as strategic signals of unwillingness to return the corrupt favour in less profitable transactions (aiming at inducing a higher transfer in the following periods) and signals of willingness to reciprocate for high transfers (aiming at receiving further high transfer in future periods in exchange for cooperation). This strategy seems to aim at the extraction of a maximum amount of cumulative bribes. In all treatments, a large fraction of non-zero transfers over all ten periods (between 36% in IDT1 and 52% in TDT2) fall into the interval for which the probability of success is significantly larger in the GD than in the ID treatments. Hence the strategies followed by groups seem to be more successful in the sense of higher reciprocal stability between briber and official than the strategies applied by individuals. We interpret this as a piece of strong evidence for the dominance of the Profit Maximizing Motive in group decision-making (**Hypothesis 3**).

5.3 Switching behaviour

So far we have analyzed revealed behaviour by considering outcomes alone. Our data on the group decision-making process allows for a more detailed analysis of the reasons and motivations underlying observed treatment differences. Since there are no significant differences in the behaviour of officials between the treatments in Stage 2 (see Section 5.1) we concentrate on O's Stage 3 behaviour.

The data on initial choices of individual officials in the GD treatments demonstrate that (at least part of) the higher levels of reciprocity within groups are due to mechanisms within the decision-making process and not based on differences in individual preferences. First, we identify situations in which officials within a group initially revealed opposing opinions on a decision, i.e. one official in the group chose to 'defect' and the other to 'cooperate' in the first step of the decision-making process. Second, we compare the (relative) numbers of successful corrupt transactions and failures following initial disagreement. For simplicity we pool cases of final disagreement and final consent against reciprocation (since both cases lead to a failed deal because of the veto power rule). Tables 3 and 4 show average percentages of corrupt success and failure conditional on initial consent (or the lack of it).

Table 3: **Success and Initial Consent, TDT1**

	Successful corruption	Failed Corruption	Total
No initial Consent	23.13%	8.75%	31.88%
Initial Consent	23.75%	44.37%	68.12%
Total	46.88%	53.12%	100%

Averages are derived from 160 transactions (16 independent groups of officials in 10 periods) in TDT1

Table 4: **Success and Initial Consent, TDT2**

	Successful corruption	Failed Corruption	Total
No initial Consent	20.42%	18.33%	38.75%
Initial Consent	21.66%	39.59%	61.25%
Total	42.08%	57.92%	100%

Averages are derived from 240 transactions (12 independent groups of officials in 10 periods) in TDT2

Assuming independence of decisions (i.e. no influence of the process on final decisions) we would expect 100% of transactions without initial consent to fail because of the veto power of the non-reciprocating official. On the contrary, we find that the final decision was made in favour of (corruption-stabilizing) reciprocity in 72.6% ($\frac{23.13}{31.88}$, TDT1 in Table 3) and 52.70% ($\frac{20.42}{38.75}$, TDT2 in Table 4) of cases in which the two officials initially disagreed. Assuming that initial decisions reflect the true underlying preferences, this means that the decision-making process alone is responsible for a large share of the treatment effects with respect to corrupt success levels. We conclude that (in both treatments) those officials who are in favour of engaging in, or maintaining, a successful corrupt relationship dominate the outcome of the decision-making process although their decision-adversaries hold veto power. We take this finding as evidence for the Persuasive Argument Theory (Pruitt 1971) which suggests that those participants (in the role of O) who provide the most valuable ideas for maximizing long term individual payoffs during the experiment (which in our case is the maintenance of the corrupt relationship through reciprocity, see Section 5.2) dominate the decisions within a group.

5.4 Content analysis

In addition to the arguments derived from the comparison of outcomes between the treatments (see Sections 5.1 and 5.2) and the analysis of choices in the different phases of the group decision-making process (see Section 5.3), we are able to get some insight into the mechanism of group decision-making by considering the content of the messages[22] exchanged during the decision-making processes of Stage 2 and Stage 3. 22 (out of 28, 16 in TDT1 and 12 in TDT2) groups exchanged electronic messages.[23] First, we separate messages and identify 132 distinct statements.[24] We allocate each statement (sent in either of the two stages) into four main categories: 'Neutral' (statements that do not contain any traceable argument, e.g. 'Hello, nice game'); 'Social' (statements including arguments against the cooperation in the corrupt transaction mentioning the negative externality, e.g. 'We have to consider the effect on the others, we should not cooperate'); 'Strategic' (arguments in favour of the stabilization of the reciprocal relationship with the objective of payoff maximization, e.g. 'Let us cooperate, otherwise we won't get any profit in the next period(s)') and 'Strategic Neg.' (arguments against cooperation in a certain period to implicitly demand larger transfers in future periods, e.g. 'Do not re-transfer, then he [the briber] will know to give more next time').[25] We add a 5^{th} category 'Social/Strategic' to account for (mostly twisted) statements that included both, other-regarding (social) and strategic (payoff maximizing) arguments. Table 7 reports the relative frequencies of statements of the respective categories subdivided by the final outcome of the respective transaction in terms of success and failure. Of all statements, only 12.2%[26] contain other-regarding arguments (Social and Social/Strategic). Their low frequency is noteworthy, and so is their lack of effectiveness (only 37.7%[27]) of transactions finally fail).

[22]We analyse all electronic chat messages exchanged by officials in the GD treatments.

[23]6 groups either did not encounter a situation of initial disagreement or ignored the possibility of writing messages.

[24]A 'conversation' between two officials may yield more than one statement since it may be split into single entries.

[25]All examples are translated (word by word) into English from the original statements in German.

[26]$8.3 + 3.9\%$, Table 5

[27]$\frac{3.1+1.5}{12.2}\%$, Table 5

Table 5: **Success and Content**

	Neutral	Social	Strategic	Strategic Neg.	Social/Strategic	Total
Success	8.3%	0.8%	27.3%	1.5%	6.8%	44.7%
Failure	31.8%	3.1%	12.1%	6.8%	1.5%	55.3%
Total	40.1%	3.9%	39.4%	8.3%	8.3%	100%

Percentages are derived from 132 statements in TDT1 and TDT2

An explanation may be that in more than 75% of all situations a social argument was followed (in the same chat conversation) by a statement arguing in favour of strategic reciprocity. 82% of these situations ended with a successful corrupt transaction. The majority (56.0%[28]) of statements contained arguments in favour of some kind of strategic reciprocity. Additional to 63 statements of positive reciprocity there were 11 separate statements arguing in favour of strategic defection aimed at extracting larger bribes in future periods. In 19 (out of all 28 or 22 relevant) groups of officials we found at least one statement in favour of strategic reciprocity (positive or negative).

The dominance of arguments in favour of payoff maximization is demonstrated not only by the relative frequency but also by the effectiveness as to corrupt success (71.5%[29] of statements including an argument for strategic (positive) reciprocity ended in a successful corrupt transaction). This provides another piece of evidence for the hypothesis that the Profit Maximizing Motive is the driving force in the decisions made in groups. Arguments that seem persuasive in the pursuit of payoff maximizing are adopted and corresponding suggestions (i.e. maintenance of strategies aiming at payoff maximizing through corrupt reciprocity) realized, while arguments in favour of social efficiency (and fairness) are neglected, since they would lead to individually costly strategies. Again the argumentation is in line with the Persuasive Argument Theory (Pruitt 1971).

We leave it to further research to separate the effect of the decision-making process from effects stemming exclusively from the nature of the exchange of arguments via electronic chat messages. For our purpose of evaluating the effectiveness of the 4EP the effort of distinguishing between those two would lead to an even more artificial setting and therefore would not help to derive conclusions.

[28] $39.4 + 8.3 + 8.3\%$, Table 5

[29] $\frac{27.3+6.8}{39.4+8.3}\%$, Table 5

6 Conclusion

The results of our experiment are interesting in two respects. First, they serve as an assessment of the usefulness of the Four Eyes Principle. Second, they provide an insight into the mechanism of group decision-making. With our experiment, using the framework of a simple 3-Stage game which is standard in the experimental corruption literature, we show that the introduction of the Four Eyes Principle, which is generally promoted as one of the most effective tools to curb corruption, can be counter-productive. We find that it increases the relative number of successful corrupt transactions as well as the amount of bribes being transferred resulting in reduced welfare (measured by the sum of participants' payoffs) and equality (measured in the distribution of payoffs across participants).

Moreover, we find two opposing effects of the introduction of the Four Eyes Principle. One, the Bribe Splitting Effect, is caused by the splitting of the transfer between two officials, which reduces the level of corruption by changing the trade-off between its costs and its benefits. The other, the Group Decision-making Effect, increases the level of corruption. We are able to separate these two effects by the use of four different treatments and show that the Group Decision-making Effect is negative and over-compensates the Bribe Splitting Effect. This leaves a negative total effect due to the introduction of the Four Eyes Principle with respect to the level of corruption and resulting social efficiency.

To explain the direction and magnitude of the Group Decision-making Effect (leading to higher conditional rates of reciprocity) we proceeded in three steps. First, the differences in revealed strategies between groups and individuals (i.e. final outcomes) can be identified by non-parametric tests and quantified by (parametric) regression analyses of conditional levels of corruption. Groups of officials reciprocate more often for high transfers and less often for low transfers than individual officials. These functional strategies lead to a higher number of successful corrupt transactions in the group decision-making treatments. Second, the analysis of behaviour within the group decision-making process provides further evidence. Contrary to predictions, in most cases initial disagreement between jointly deciding officials leads to a successful corrupt transaction despite the veto-power of non-corrupt officials. Third, we analyze the content of electronic chat-messages, exchanged during the decision-making process. Arguments in favour of strategic reciprocity (i.e. initiating or maintaining only corrupt transactions that yield a large payoff through high transfers) dominate the decision making-process not only quantitatively but also in terms of effectiveness (outcomes). The results of this 3-step analysis suggest the dominance of the 'Profit Maximizing Motive' in the group decision-making process (Kocher and Sutter 2007). This is in line with the Persuasive

Argument Theory (Pruitt 1971).

So, policies that prescribe group decision-making should be restricted to situations for which the pursuit of maintaining reciprocal and payoff maximizing strategies are in line with the policies' objectives. Our results cast serious doubt on the usefulness of the Four Eyes Principle for situations where this condition does not apply. Looking into the black box of the mechanism that underlies group decision-making through the analysis of processes in combination with content analyses within controlled laboratory experiments may help to interpret behavioural patterns and to discover determinants of situations where the strategic use of group decision-making might reduce social inefficiencies. Future research should be directed at the theoretical foundation of the mechanisms found within group decision-making.

Appendix

Appendix A

Equilibrium in the 3-Stage Game

Proof by (backward Induction).

Denote by $I_{i,n}$ the information set in stage i ($i \epsilon \{1,2,3\}$) of period n ($n \epsilon \{1,2,...,10\}$). Let $p(I_{i,n})$ be the probability of reaching the respective stage and $q(\text{`}s'|I_{i,n})$ the conditional probability of the relevant agent choosing action 's' once reached Stage (i,n). An information set contains all relevant information about ego's and alter's behaviour up to the respective stage. Furthermore let $PO(I_{i,n})$ be the (sum of) payoff(s) gained up to the arrival of stage (i,n).

First we show that there cannot be an equilibrium in which O chooses 'cooperate' in Stage 3 of the last (10^{th}) period. Consider a Strategy-Set $EQU1 = [s_{1,1}, s_{2,1}, s_{3,1}, s_{1,2}..., s_{3,10}]$ in which the third stage of period 10 is reached with some probability ($p(I_{3,10}) > 0$) and O cooperates with some probability ($q(\text{`}cooperate\text{'}|I_{3,10}) > 0$). Compare the payoff, resulting from the realization of Strategy-Set $EQU1$ ($PO(EQU1)$) to the one of an alternative Strategy-Set $EQU1_{new}$ which consists of the same strategies up to $I_{3,10}$ but for which $q(\text{`}cooperate\text{'}|I_{3,10}) = 0$ yielding payoff $PO(EQU1_{new})$.

Since the payoff for period 10 is larger for $EQU1_{new}$, since $8 + 3 * b < 12 + 3 * b$, $EQU1$ cannot constitute a Sub-game perfect Nash equilibrium. Second we show that, in the last (10^{th}) period, B will never choose any Strategy-Set that includes the action '$b > 0$' in Stage 1. Consider again a Strategy-Set $EQU2 = [s_{1,1}, s_{2,1}, s_{3,1}, s_{1,2}, ..., s_{3,10}]$ in which $p(I_{1,10}) > 0$, $q(\text{`}cooperate\text{'}|I_{3,10}) = 0$ and $q(\text{`}b > 0\text{'}|I_{1,10}) > 0$.[30] Again compare $PO(EQU2)$ to $PO(EQU2_{new})$, the payoff of a Strategy-Set that differs from the former only in $q(\text{`}b > 0\text{'}|I_{1,10}) = 0$. Since $12 - b \leq 12$, payoff $PO(EQU2)$ must be smaller than $PO(EQU2_{new})$ so that $EQU2$ cannot constitute an equilibrium. Hence only a Strategy-Set featuring $[s_{1,1}, ..., s_{9,1}, \text{`}b = 0\text{'}, \text{`}accept\text{'}/\text{`}reject\text{'}, \text{`}defect\text{'}]$ can characterize an equilibrium.

Consider now a period-set $PS = \{k, ..., 10\}$ of (the last 10-k) consecutive periods for which the above stated last period's equilibrium Strategy-Set is played. Assume $q(\text{`}cooperate\text{'}|I_{3,k-1}) > 0$ for the period $k-1$. By the same line of arguments as for the last (10^{th}) period we can easily repeat the task up to the point of excluding all strategy

[30] Given that $q(\text{`}cooperate\text{'}|I_{3,10}) = 0$ must be satisfied, O will never choose 'reject' in Stage 2 for '$b > 0$' and is indifferent between 'reject' and 'accept' if '$b = 0$', see Appendix 1C in Chapter 1.

sets that do not exhibit the strategy characteristics of the (Stage Game) equilibrium in the 10^{th} period [‘$b = 0$’, ‘$accept$’/‘$reject$’, ‘$defect$’]. Letting k decrease from 9 down to 1, it is obvious that the Stage Game Nash Equilibrium remains the only Sub Game perfect Nash Equilibrium in the (finitely) repeated game.

Appendix B

Extensive forms of games in all treatments

In all treatments except TDT2 both, O and B, decide once in every period. In TDT2 only B decides once per period while each O decides twice.

Figure 5: Extensive forms of TDT1 and IDT1

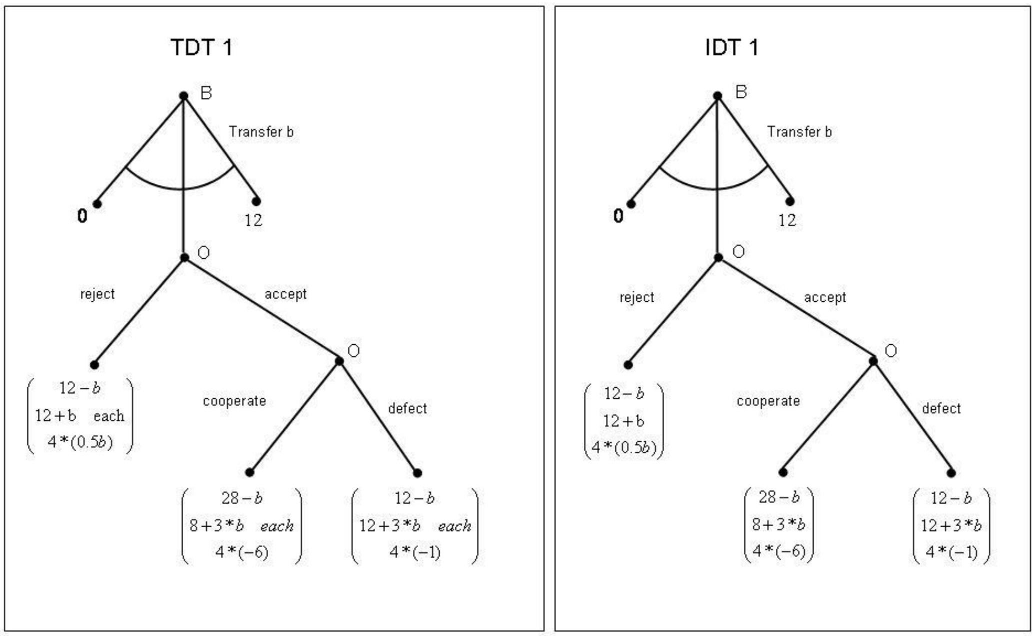


Figure 6: Extensive forms of TDT2 and IDT2

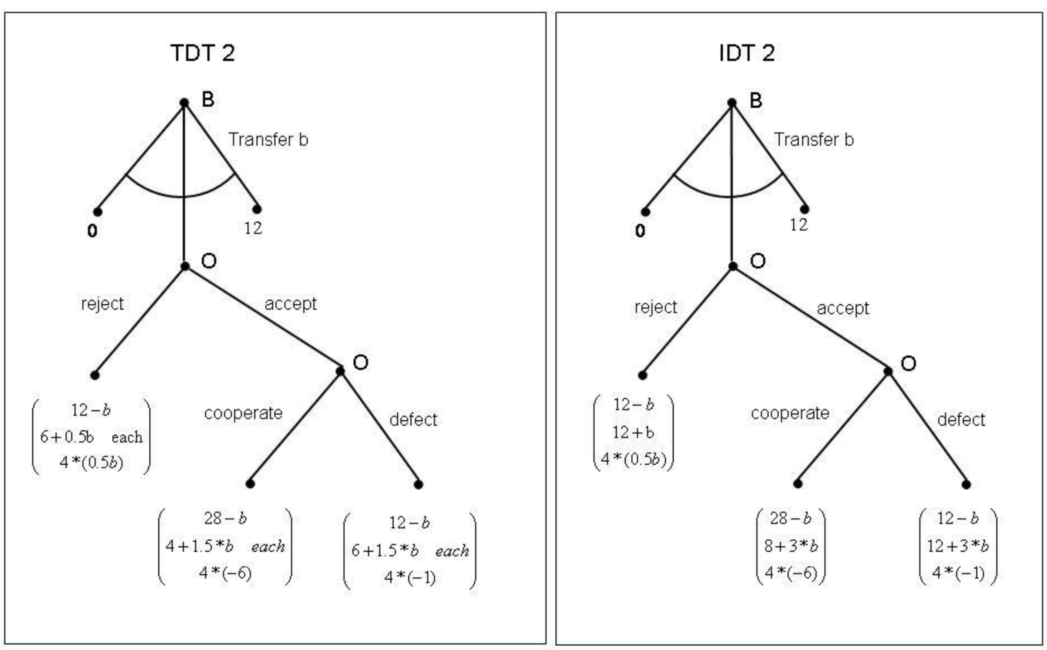


Table 6: **Model M1, Random effects estimation**

	Dependent variable: PP			
	Linear (OLS)		Tobit	
	Coefficient	Stand. error	Coefficient	Stand. error
Constant	22.2525***	1.02	22.1728***	1.92
b	1.8964***	0.28	1.8776***	0.35
b^2	−0.2727	-0.29	−0.2734	-0.32
D_{IDT2}	−0.1221	0.19	−0.1456	0.19
D_{TDT1}	−6.3437***	0.84	−6.3332***	1.14
D_{TDT2}	−3.0786***	0.46	−3.1232***	0.65
D_{IDT2}*b	0.1321	0.17	0.1432	0.19
D_{TDT1}*b	2.0542***	0.34	2.0318***	0.42
D_{TDT2}*b	0.7415***	0.19	0.7332***	0.21
D^2_{IDT2}*b	−0.0208	0.14	−0.0328	0.10
D^2_{TDT1}*b	−0.2528	0.21	−0.2421	0.28
D^2_{TDT2}*b	0.4451	0.31	0.4721	0.52
	overall $R^2 = 0.54$		-	

Number of periods: 10, Number of observation: 88, Number of clusters: 64 (heteroskedasticity) robust standard errors in OLS

Table 7: **Model M2**

	Dependent variable: Payoff			
	OLS		Tobit	
	Coefficient	Stand error	Coefficient	Stand error
Constant	225.2521***	54.92	224.5843***	44.47
N	15.8923***	2.48	15.5952***	3.11
D_{IDT2}	−7.1245	10.29	−7.0319	8.39
D_{TDT1}	−156.3421***	42.87	−157.2822***	44.27
D_{TDT2}	−133.17***	39.72	−134.0318***	41.90
D_{IDT2}*N	−0.8177	1.17	−0.7849	1.06
D_{TDT1}*N	3.1425***	0.73	3.1929***	0.75
D_{TDT2}*N	5.0308***	0.95	5.1121***	0.78
	adjusted $R^2 = 0.43$		-	

*** denotes significance at the 1%-level; Number of observations: 76; Number of clusters: 64 (heteroskedasticity) robust standard errors

Table 8: **Model M4**

	Dependent variable: SC					
	transfer level					
	0	1	2	3	4	5
b	0.026***	0.034***	0.045***	0.055***	0.065***	0.073***
$D_{IDT2}*b$	−0.002	-0.002	-0.003	-0.004	-0.004	-0.005
$D_{TDT1}*b$	0.018***	0.025***	0.032***	0.039***	0.046***	0.052***
$D_{TDT2}*b$	0.015***	0.020***	0.026***	0.032***	0.037***	0.043***
$\hat{SC}_{IDT1}$	0.093	0.119	0.154	0.198	0.254	0.319
$\hat{SC}_{IDT2}$	0.092	0.116	0.147	0.192	0.241	0.303
$\hat{SC}_{TDT1}$	0.013	0.044	0.091	0.172	0.239	0.352
$\hat{SC}_{TDT2}$	0.009	0.034	0.073	0.146	0.195	0.318
	transfer level					
	6	7	8	9	10	11
b	0.080***	0.083***	0.082***	0.078***	0.071***	0.062***
$D_{IDT2}*b$	−0.005	-0.006	-0.006	-0.005	-0.005	-0.005
$D_{TDT1}*b$	0.057***	0.059***	0.058***	0.056***	0.051***	0.044***
$D_{TDT2}*b$	0.047***	0.048***	0.048***	0.045***	0.041***	0.036***
$\hat{SC}_{IDT1}$	0.392	0.471	0.555	0.638	0.715	0.785
$\hat{SC}_{IDT2}$	0.370	0.445	0.523	0.598	0.672	0.737
$\hat{SC}_{TDT1}$	0.510	0.627	0.723	0.889	0.921	0.998
$\hat{SC}_{TDT2}$	0.471	0.542	0.698	0.802	0.897	0.967
	Number of subjects: 96					
	Number of units: 64					

*** denotes significance at the 1%-level

Marginal effects are calculated at the respective values of transfer and at the means of the remaining independent variables

$\hat{SC}$ denotes the estimates for the success probabilities conditional on the respective transfer levels.

Appendix C: Instructions from TDT2 (translated from German)

Thank you very much for your appearance. In the next 90 minutes you will take part in an experiment in the laboratory of MELESSA. If you read the following instructions carefully, you can (depending on your decisions) earn money, additional to the show-up fee of 4 Euros. Additional to the money you can earn for yourself, you will affect the amount of donation to the public aid organization 'Doctors without Borders'. The money you will earn during the experiment will be added to the show-up fee and paid out in cash at the end of the experiment. The money that is going to be donated will be transferred to the donations account of 'Doctors without Borders'.
During the experiment you are not allowed to communicate with the other participants. If you have questions, please approach one of the experimenters by raising your hand. In the case of violation of this rule we have to exclude you from any payments.
During the experiment we will refer to Experimental Monetary Units (EMU) instead of Euros. Your income will be calculated in EMU. In the end of the experiment the total amount will be exchanged in Euros.
The exchange rate is **1 EMU = 5 Eurocents**.
All 24 participants are assigned to **groups of four**. Neither the experimenters nor the other participants know which group you are in. Your decisions remain completely anonymous.

The Decision Situation

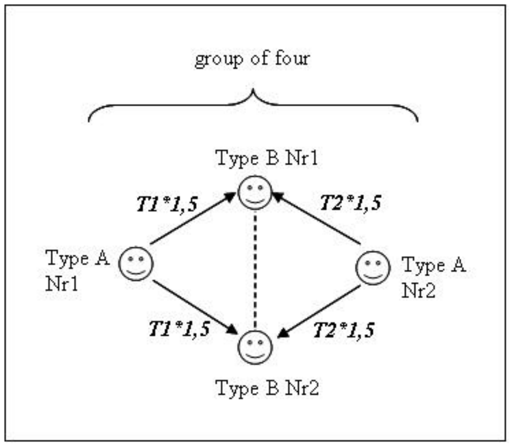
There are two types in this experiment: type A and type B. The types play different roles and make decisions that affect their own income, the income of the other participants of the experiment and the amount of donation transferred to the organization 'Doctors without Borders'. The type of a participant is allocated randomly.
A group of four consists of two type A and two type B participants who stay together for the entire experiment.
The experiment has 10 periods.

Procedure:

All of the 10 periods consist of at most 3 Stages.

Stage 1

In the first Stage, **every participant of type A** (type A Nr 1 and type A Nr 2) decides on the size of their transfer (**T1** denotes type A Nr 1's transfer and **T2** denotes type A Nr 2's transfer) which has to lie between 0 and 12 EMU.
Next, the amount of the transfers is **tripled** and **then split** equally between the two type B participants (type B Nr 1 **and** type B Nr 2) of the group of four. If T1 is for example 6 EMU, type B Nr 1 receives 9 EMU ($0.5 * 6 * 3$ EMU) and type B Nr 2 receives 9 EMU.
Hence there are 2 situations per group in any period:
Situation 1: Type A Nr1 transfers T1 to the two type B participants (where T1 is **first tripled and then shared**)
Situation 2: Type A Nr2 transfers T2 to the two type B participants (where T2 is **first tripled and then shared**)



Stage 2

In Stage 2 the two type B participants decide jointly on how to react on the transfer of the respective type A participant. They have (in both situations) two alternatives.
1^{st} Alternative: Both decide (for a specific transfer, e.g. T1) jointly for 'keep': In this case Stage 3 is entered
2^{nd} Alternative: One or both decide in favour of 'distribute': In this case, the respective type A participant (e.g. type A Nr 1) does not get a bonus (and receives only 12 - T1 EMU). The type B participants both get 6 EMU plus half of the value of the transfer ($6 + 0.5 * T1$ EMU). Moreover, the amount of $2 * T1 + 24$ EMU is transferred as a donation to the organization 'Doctors without Borders'.

A joint decision between the two subjects is found as follows.
First, **each** of the two type B participants decides **individually** whether to 'keep' or to 'distribute' the particular transfer.
If the decision is not unanimous (one type B participant wants to 'keep' and the other wants to 'distribute' the transfer), the decision of the fellow participant appears on his or her own screen.
Next, the participants decide **once again separately**. If there is still no agreement, the two type B participants can exchange messages via an electronic 'chat' (see explanation below) for one minute. After this the participants decide for the **last time**.
Note that only if both type B participants decide in favour of 'keep' the third Stage is actually reached.
Since there are two type A participants in every group of four (type A Nr 1 and type A Nr 2), each of the type B participants has to decide (jointly with the other type B participants) in two situations: once for T1 and once for T2.

Stage 3

In Stage 3 (which is only reached if both type B subjects have chosen 'keep') the two type B participants decide **again jointly** whether to initiate a re-transfer or not.
Again, both type B subjects decide **separately first**.
If the decision is not unanimous (one type B participant wants to initiate the re-transfer and the other does not), the decision of the other participant is shown on the screen. Then the participants can **decide again separately**. If there is still no consent, the participants enter again a 'chat' in which they can exchange electronic messages for one minute. After this, there is a final decision.

1. **Case**: **Both** type B participants decide **in favour of** a re-transfer. Both carry the costs of 2 EMU each (independent of the amount of the respective transfer). They both get 6 EMU plus one and a half times the value of the transfer, less the costs of 2 ($6 + 1.5 * T1 - 2$ EMU). The respective type A participant (type A Nr 1) receives a Bonus of 16 EMU in addition to the 12 EMU of initial endowment ($16 + 12 - T1$). In this case there is no donation to the organization 'Doctors without Borders'.
2. **Case**: **One or both** type B participants decide **against** a re-transfer. In this case, there are no personal costs for the two type B participants (they get $6 + 1.5 * T1$ each), the respective type A participant does not receive a bonus (and gets $12 - T1$), and the donation to the organization is 20 EMU.

In the end, all participants are shown their personal income in the period. Please note, that the type A participants can thereby reconstruct whether or not the type B participants chose for or against the re-transfer.
These (maximal) 3 stages are repeated 10 times (10 periods). Since the members of groups stay together, participants *always* **interact with the same persons** in the same roles for the **entire experiment**. (Type A Nr 1 remains type A Nr 1. type A Nr 2 remains type A Nr 2 etc.)

Chat:

Type B subjects potentially have the possibility to communicate via real time **electronic messaging** (Chat) with their fellow type B subject to agree on a joint decision (e.g. 'keep' or 'distribute') in Stage 2 and Stage 3.
The content of the communication is generally free to choose but there are some restrictions. You are not allowed to make statements about personal characteristics such as your name, age, address, gender, subject of study or any information that might lead to your identification. Moreover, strong language is strictly forbidden. Anyone who violates these rules of communication will be automatically expelled from the experiment and will not get any payments **for the entire experiment**.
Each participant in the chat can send as many messages to the other participant as he wishes or is able to send within the time limit of one minute.
Every message appears automatically on the screens of both type B participants of a group of four but cannot be seen by any other participant of the experiment.

Payoff table

The following table shows the kind of consequences the decisions of the participants lead to - in terms of their own payoff, the payoff of the other participants and the organization 'Doctors without Borders' (Example for T1).
The following table can be read as follows. Generally we start from the top and go down cell by cell. If a participant chooses a certain alternative, only those cells that lie directly beneath it are relevant for the next period.

The payoff table is analogous for situations in which T2 (Transfer of Type A Nr2) is relevant.

Note that each **type B participant** receives **two** payments because two situations are relevant for each of them, one with type A Nr 1 (T1 is relevant) and one with type A Nr 2 (T2 is relevant). These two are added up for any period.
For the **type A participants** only **one** situation per period is relevant so that there is only one payment per period.

1. Stage	Type A Nr1 chooses ***T1***		
2. Stage	At least one Type B participant (Type B Nr1 **or** Type B Nr2 **or both)** decide in favour of ***`distribute´*** for ***T1***	**Both** Type B participants (Type B Nr1 **and** Type B Nr2) decide in favour of ***`keep´*** for ***T1***	
3. Stage	Stage 3 is not reached	Type B Nr1 or Type B Nr2 or both decide against a **re-transfer** to Type A Nr1	**Both** (Type B Nr1 and Type B Nr2) decide in favour of a **re-transfer** to Type A Nr1
Payoff **Type A Nr1**	**12 –T1**	**12 –T1**	12 + 16–T1 = **28 –T1**
Payoff **Type B Nr1/Type B Nr2**	**6 + 0,5*T1**	**6 + 1,5*T1**	6 - 2 + 1,5*T1 = **4 + 1,5*T1**
Donation to `Doctors without Borders´	**2*T1+ 24**	**20**	**0**

Each participant gets information at the end of each period about his own personal payoff. Type A subject can infer whether type B subjects have chosen to initiate a re-transfer or not.
Note that the sum of payments (exchanged in Euros) to the organization is actually donated to 'Doctors without Borders'.

Timing

Stage 1: Type A Nr1 chooses T1 and Type A Nr2 chooses T2
Stage 2: Type B Nr1 and Type B Nr2 decide (T1 and T2) each time jointly about 'keep' or 'distribute'
Stage 3:
Situation (for T1): Only reached if in this situation, both Type B participants chose 'keep' in Stage 2. In this case, both Type B participants decide jointly whether to initiate a re-transfer or not.
Situation (for T2): is analogous, only for T2.

At the end of each of the ten periods, each participant gets information about his/her own payoff in the respective period. At the end of the last (10^{th}) period, participants get to know their final income and their payment in Euros.

The following **control questions** will help you to get a better understanding of the situation. All the necessary information can be found in the payment table.
Please answer all the control questions and raise your hand when you have finished. An experimenter will come to your place to check your solutions.

Question 1

Assume that **you are type A Nr1** and you chose a transfer of **4 EMU** (T1). The other participant of type A (type A Nr2) has chosen a transfer (T2) of **10 EMU**.
Situation 1 (T1): One of the participants of type B in your group of four (type B Nr1) decides to **'distribute'** your transfer (T1). The other participant of type B (type B Nr2) wants to **'keep'** your transfer (T1). (Therefore Stage 3 is not reached.)
Situation 2 (T2): Both type B participants chose to **'keep'** the transfer of type B Nr 2 (T2) in Stage 2 and decide **against** a re-transfer in Stage 3.

a) What is the payoff of type B Nr1 in the situation (1) with type A Nr1 (you)?
Your answer:___________________________
b) What is the payoff of type B Nr2 in the situation (1) with type A Nr1 (you)?
Your answer:___________________________
c) What is your (type A Nr1) **total** payoff in this period?
Your answer:___________________________
d) What is the total payoff of type B Nr1 for all situations relevant to him/her?
Your answer:___________________________
e) What is the total payoff of type B Nr2 for all situations relevant to him/her?
Your answer:___________________________
f) What is the total payoff of type B Nr2 in this period?
Your answer:___________________________
g) What is the amount of donation to 'Doctors without Borders' caused by the situation relevant to you (type A Nr1)?
Your answer:___________________________
h) What is the amount of donation to 'Doctors without Borders' caused by the situation relevant to type A Nr2?
Your answer:___________________________
i) What is the total amount of donation to 'Doctors without Borders' in this period?
Your answer:___________________________
j) What is the total amount of payoff generated by the decisions of your group of four?
Your answer:___________________________

Question 2

Assume that you (type A Nr1) and type A Nr2 have both chosen a transfer of **0** (T1 is 0 EMU and T2 is 0 EMU). Neither participant of type B (neither type B Nr1 nor type B Nr2) wants to 'keep' any of the two transfers in Stage 2.
a) What is your (type A Nr1) total payoff in this period?
Your answer:___________________________
b) What is the total payoff of type A Nr2 in this period?
Your answer:___________________________
c) What is the total payoff of type B Nr1 in this period?
Your answer:___________________________
d) What is the total payoff of type B Nr2 for all situations relevant to him/her?
Your answer:___________________________
e) What is the total amount of donation to 'Doctors without Borders' in this period?
Your answer:___________________________
f) What is the total amount of payoff generated by the decisions of your group of four?
Your answer:___________________________

Question 3

Assume that you (type A Nr1) have chosen a transfer of **5 EMU** (T1 is 5 EMU) and type A Nr2 has also chosen a transfer of **5 EMU** (T2 is 5 EMU). **Both** participants of type B (type B Nr1 and type B Nr2)

decide to **'keep'** the transfer and initiate a re-transfer in Stage 3.
a) What is your (type A Nr1) total payoff in this period?
Your answer:__________________________
b) What is the total payoff of type A Nr2 in this period?
Your answer:__________________________
c) What is the total payoff of type B Nr1 in this period?
Your answer:__________________________
d) What is the total payoff of type B Nr2 for all situations relevant to him/her?
Your answer:__________________________
e) What is the total amount of donation to 'Doctors without Borders' in this period?
Your answer:__________________________
f) What is the total amount of payoff generated by the decisions of your group of four?
Your answer:__________________________

References

Abbink, K. (2004). 'Staff Rotation as an Anti−corruption Policy: an Experimental Study', *European Journal of Political Economy*, VOL 20, pp. 887−906

Abbink, K., H. Hennig-Schmidt (2006). 'Neutral versus Loaded Instructions in a Bribery Experiment', *Experimental Economics*, VOL 9/2, pp. 103−21

Abbink, K., B. Irlenbusch, E. Renner (2002). 'An Experimental Bribery Game', *Journal of Law, Economics and Organization*, VOL 18/2, pp. 428−54

Anderson, L., C. Holt (1997). 'Information Cascades in the Laboratory', *American Economic Review*, VOL 87, pp. 847−62

Becker, G. (1968). 'Crime and Punishment: an Economic App.roach', *Journal of Political Economy*, VOL 78, pp. 169−217

Berg, J., J. Dickhaut, K. McCabe (1995). 'Trust, Reciprocity, and Social History', *Games and Economic Behavior*, VOL 10, pp. 122−42

Bishop G., D. Myers (1974). 'Informational Influences in Group Discussion', *Organizational Behavior and Human Performance*, VOL 12, pp. 92−104

Blinder, A., J. Morgan (2005). 'Are Two Heads Better than One? An Experimental Analysis of Group vs. Individual Decision Making', *Journal of Money Credit and Banking*, VOL 37, pp. 789−812

Bone J., J. Hey, J. Suckling (1999). 'Are Groups more (or less) Consistent than Individuals?', *Journal of Risk and Uncertainty*, VOL 8, pp. 63−81

Bornstein G., I. Yaniv (1998). 'Individual and Group Behavior the Ultimatum Game: are Groups more 'Rational' Players?', *Experimental Economics*, VOL 1, pp. 101−08

Burnstein E., A. Vinokur, Y. Trope (1973). 'Interpersonal Comparison versus Persuasive Argument: a more Direct Test of Alternative Explanations for Group-induced Shifts in Individual Choices', *Journal of Social Psychology*, VOL 9, pp. 236−45

Cason, T., V. Mui (1997). 'A Laboratory Study of Group Polarization in the Team Dictator Game', *Economic Journal*, VOL 107, pp. 1465−83

Chalos P., S. Pickard (1985). 'Information Choice and Cue use: An experiment in Group Information Processing', *Journal of Applied Psychology*, VOL 70, pp. 634−41

Charness, G., M. Rabin (2002). 'Understanding Social Preferences with Simple Tests', *Quarterly Journal of Economics*, VOL 117, pp. 817–69.

Fehr, E., G. Kirchsteiger, A. Riedl (1993). 'Does Fairness Prevent Market Clearing? An Experimental Investigation', *Quarterly Journal of Economics*, VOL 108, pp. 437–59

Fehr, E., K. Schmidt (1999). 'Theory of Fairness, Competition and Cooperation', *Quarterly Journal of Economics*, VOL 114, pp. 817–68

Fehr, E., K. Schmidt (2003). 'Theories of Fairness and Reciprocity–Evidence and Economic Applications', *in Dewatripont, M., L.P. Hansen, S.J. Turnovsky, 'Advances in economics and econometrics'*, 8th World Congress, Economic Society Monographs, pp. 208–57

Fischbacher, U. (2007). 'Z-Tree: Zurich Toolbox for Ready-made Economic Experiments', *Experimental Economics*, VOL 10/2, pp. 171–8

Gächter, S. (2007). 'Conditional Cooperation: Behavioral Regularities from the Lab and the Field and their Policy Implications', *in Frey, B.S., A. Stutzer, 'Economics and Psychology. A Promising New Cross-Disciplinary Field'*, MIT Press, Cambridge, MA, pp. 19–50

Gneezy, U., J. A. List (2006). 'Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments', *Econometrica*, VOL 74, pp. 1365–84

Heckman, J. (1979). 'Sample Selection Bias as a Specification Error', *Econometrica*, VOL 47, pp. 153–61

Klitgaard, R. (1988). 'Controlling Corruption', *University of California Press*

Kocher M., M. Sutter (2005). 'The Decision Maker Matters. Individual versus Team Behavior in Experimental Beauty-contest Games', *Economic Journal*, VOL 115, pp. 200–23

Kocher M., M. Sutter (2007). 'Individual versus Group Behavior and the Role of the Decision making Procedure in Gift–Exchange Experiments', *Empirica*, VOL 34, pp. 63–88

Kugler, T., G. Bornstein, M. Kocher, M. Sutter (2007). 'Trust Between Individuals and Groups: Groups are Less Trusting than Individuals but just as Trustworthy', *Journal of Economic Psychology*, VOL 28, pp. 646–57

Lambsdorff, J. Graf (2007). 'The Institutional Economics of Corruption and Reform, Theory, Evidence and Policy', *Cambridge University Press*

Lambsdorff, J. Graf, B. Frank (2007). 'Corrupt Reciprocity – an Experiment', *Diskussionsbeitrag Nr. V–51–07, Volkswirtschaftliche Reihe*, ISSN 1435–3520

Lambsdorff, J. Graf, B. Frank (2010). 'Bribing versus Gift-giving – An Experiment', *Journal of Economic Psychology*, VOL 31/3, pp. 347–57

Levine, J., R. Moreland (1998). 'Small groups', in *'Handbook of social psychology'*, VOL 2/4, pp. 415–69

OECD (2010). 'Guidelines for Managing Conflict of Interest in the Public Service', *http://www.oecd.org/ dataoecd/ 31/15/ 36587312.pdf*

Pereira, P.T., N. Silva, J. Andrade e Silva (2006). 'Positive and Negative Reciprocity in the Labor Market', *Journal of Economic Behavior & Organization*, VOL 59, pp. 406–22

Pörting, P., W. Vahlenkamp (1998). 'Internal Strategies Against Corruption: Guidelines for Preventing and Combatting Corruption in Police Authorities', *Crime Law and Social Change*, VOL 29, pp. 225–49

Pruitt, D.G. (1971). 'Choice Shifts in Group Discussion: An Introductory Review', *Journal of Personality and Social Psychology*, VOL 20/3, pp. 339–60

Rieger, H.-J. (2005). 'Prevention – a Key Factor in Fighting Corruption', *United Nations Public Administration Network*

Rose-Ackerman, S. v. (1999). 'Corruption and Government, Causes, Consequences and Reform', in *Lambsdorff, J. Graf 'The New Institutional Economics of Corruption'*, Cambridge University Press 1999

Schulze, G., B. Frank (2003). 'Deterrence versus Intrinsic Motivation: Experimental Evidence on the Determinants of Corruptibility', *Economics of Governance*, VOL 4/2, pp. 143–60

Tajfel, H., J.C. Turner (1986). 'The Social Identity Theory of Intergroup Behavior', *in Worchel, S., W.G. Austin 'Psychology of intergroup relations'*, Chicago, pp. 7–24

Wiehen, M. (2005). 'Global Corruption Report 2005', *Transparency International*, Pluto Press, London and Ann Harbor