Social networks and intergroup conflict

Takács, Károly

*"…in group relations, some degree of prejudice-hostility always appears when there is the combination of **visibility** and **competition**."*

Robin M. Williams, Jr.: The Reduction of Intergroup Tensions. A Survey of Research on Problems of Ethnic, Racial, and Religious Group Relations (1947: 54)

CHAPTER 3

SOCIAL CONTROL IN INTERGROUP COMPETITION

Experimental evidence

3.1 Introduction

Single-shot social dilemma experiments often find unexpectedly high cooperation rates. Even in strictly impersonal settings, in which no communication is allowed and subjects are completely strangers, many people tend to act against their own interests and make sacrifices for the collectivity. In a competition situation with another group, it is even more likely that subjects put their direct monetary incentives aside in favor of contributing to the public good provision (Bornstein and Ben-Yossef, 1994). In intergroup competition experiments (e.g., Bornstein, Erev, and Rosen, 1990), subjects were not allowed to redistribute payoffs, to invent binding commitments, or to apply coercion. High level of cooperation was nevertheless achieved.

Why do people tend to act in favor of their groups if it is not in their interest to do so? This question is particularly challenging to answer when widespread cooperation within competing groups has lethal consequences for members of both groups. A harmful outcome provides just another incentive against contribution, but people still not refrain from making sacrifices of this kind for their group. Such action is difficult to explain assuming perfect individual rationality. In this chapter, we do not discuss limitations of rationality, but we raise doubts about whether the monetary payoff structure of experimental games fully describes the incentives of players in the laboratory. We claim that there are also other substantial utility concerns. We will concentrate on the role of incentives that stem from *structural embeddedness*. We try to demonstrate that elementary structures might be present in the laboratory and can make a significant difference to contribution decisions, even when subjects do not know each other and are not allowed to communicate.

In this way, this is the very first attempt to incorporate structural embeddedness in the experimental analysis of intergroup competition. As a model of intergroup relations, we will use the structurally embedded Intergroup Public Goods (IPG) game that was discussed in Chapter 2. We developed this model in order to explain how structural embeddedness influences the likelihood of intergroup conflict. The theoretical results we derived for this question will guide us as hypotheses in the experiments. Among effects of structural embeddedness this study is particularly interested in the effect of *segregation* on intergroup conflict. In the theoretical analysis of Chapter 2 we found that in general the effect of segregation on the likelihood of intergroup conflict can be characterized by an S-shape function. This implies that segregation is likely to promote intergroup conflict, but in certain ranges of segregation, an additional change does not result in an increase in the likelihood of conflict. These results directly lead to one of the main research questions we would like to answer in the experiments. This regards whether the likelihood of intergroup conflict is higher when group members are arranged in a segregated pattern in the experiments.

As an underlying mechanism, in Chapter 2 we analyzed the impact of *social control* on individual decisions. In previous experiments, there has been little focus on structural embeddedness and social control. There are explanations of seemingly irrational cooperation that concentrate on group identity, on fairness, or on other related concepts (social commitment, conformity), but we do not exactly know how these processes work in practice. There are some indications of the presence of social control in earlier social dilemma experiments (van de Kragt, Dawes, and Orbell, 1988; Rapoport, Bornstein, and Erev, 1989), but without clear conceptualization of these effects.

As in Chapter 2, in the experiments we focus on three fundamental forms of social control. The first is *behavioral confirmation* that expresses the subject's desire to conform to the expected behavior of other individuals. The second is the distribution of *social selective incentives,* such as prestige or respect. The third form is present in network relations between members of the opposite groups and is referred to as a *traitor reward.* It is a social selective incentive that punishes contribution and rewards defection. For the derivation of hypotheses about these social control mechanisms, results from Chapter 2 give directives. These results show that the larger the size of selective incentives compared to behavioral confirmation, the stronger the effect of segregation on intergroup conflict. In the experiments, we explore in which forms social control influences individual behavior and whether the segregation effect on intergroup conflict is indeed stronger under normative pressure than under confirmation pressure.

For the conceptualization of social control in the laboratory, we have to consider that social control might become *internalized* (Scott, 1971; Coleman, 1990: 293). The internalization of *contribution norms* creates a cognitive reward for cooperation. This motive, that can be summarized as "I do the right thing for the group", is relevant even if individual decisions are completely confidential (Opp, 1989). Similar explanations were provided for surprisingly high contribution rates in social dilemmas and for high participation rates in collective action situations from various theoretical perspectives. Such motivations have already been recognized by Smith (1976[1759]) and were as a "universal fellow feeling". The interpersonal character of these intrinsic motivations were acknowledged by Sen (1977), who called them "social commitment", and by Elster (1990), who specified these behavioral incentives as a "sense of duty". They argued that social commitment and respectively the sense of duty is obviously stronger in the direction of somehow related individuals (cf. Caporael et al., 1989). Others attributed the selective character of internalized contribution norms to internalized values of altruistic or non-selfish behavior (Kenrick, Baumann, and Cialdini, 1979; Cialdini et al., 1987).

Some concepts that have appeared in the literature as possible solutions for social dilemmas show similarity with *behavioral confirmation*. *Reciprocity* can be an efficient behavioral pattern in repeated social dilemmas (Axelrod, 1984), but it can also be present in single-shot situations in a cognitive form. For instance, individuals might believe that others are good contributors and they might believe that others have similar beliefs. Furthermore, they might also think that there is a mutual consensus about considerations of reciprocal fairness (cf. Camerer, 1997). Besides reciprocity, another mechanism that shows similarity with behavioral confirmation is *imitation*. Even if others are not able to monitor individual choice, people tend and like to behave as others or imitate behavioral samples (Asch, 1956; Dawkins, 1976; Pingle, 1995). In classical sociology, Weber (1978[1913]: Chapter I.4, 29) specified “custom” similar to internalized behavioral confirmation: “As distinguished from both ‘convention’ and ‘law’, ‘custom’ refers to rules devoid of any external sanction. The actor conforms with them of his own free will…” In a certain sense, fashion works in a similar way, as individuals gain enjoyment by confirming the behavior of relevant others (e.g., Richardson, 1948a; 1948b).

Traitor rewards are also likely to be internalized as a fear of local conflict and benefit for local harmony. Betraying someone else elicits an unpleasant feeling of guilt; therefore people try to avoid this (e.g., Poundstone, 1992: 223).

We believe that these forms of social control influence individual contribution decisions in intergroup related collective action. As an aggregated consequence of dyadic social control, the network structure of individual relations influences the likelihood of intergroup conflict. Dense in-group relations and scarce out-group relations are correlated with extensive distribution of social selective incentives between fellows and limited realization of traitor rewards. Hence, segregation supports the emergence of harmful intergroup clashes. The key elements of the explanatory mechanism are the fundamental forms of social control. In this chapter, we will examine what types of direct and internalized social control influence the decision of subjects to contribute or not to group collective action in controlled experimental conditions. In short, we are interested in what forms of social control back the segregation effect on intergroup conflict.

To test the presence of different forms of social control and the segregation effect on intergroup conflict, we introduced a *new experimental design*. Seating patterns were varied and visibility conditions were manipulated in order to detect forms of social control that are activated conditional on the neighborhood composition. We implemented a full segregation, a complete mixture, and a medium segregation condition. Additionally, in the control condition, subjects made decisions in isolation. As an aggregated consequence of local social control we derived a prediction that intergroup conflict will be more likely in segregated seating configurations. This prediction was tested in the experiments.

Subjects were unknown to each other. They were not able to communicate in any form and there was no way to learn of the decisions of others. Consequently, there were no real network relations between the subjects. For testing the presence of internalized social control, we introduced *minimal contact* between neighboring subjects. Minimal contact meant that subjects were able to see each other and they were able to identify the group membership of each other. We were interested in whether this minimal contact is sufficient to activate internalized forms of social control in the experiment.

Experimental evidence shows that face-to-face contact facilitates cooperation in conflict situations (Drolet and Morris, 2000). Previously, this finding was explained by the social psychological process of rapport that is conceptualized as a “state of mutual positivity and interest that arises through the convergence of nonverbal expressive behavior in an interaction” (Drolet and Morris, 2000: 27; Tickle-Degnen and Rosenthal, 1990). There is no doubt that when subjects are able to communicate with nonverbal signs or are able to send emotional signals, they influence the behavior of each other in the social dilemma task. However, we would like to emphasize that minimal contact has an additional effect that is due to the activation of internalized social control. To show this, extensive nonverbal communication was disallowed in the experiment and decision times were restricted to avoid initiations of signaling.

Our experiments provide a test whether we can partly explain a higher level of contribution in the IPG game by internalization of social control (behavioral confirmation, selective incentives, and traitor rewards) under strictly impersonal experimental conditions with only a minimal contact between subjects. In later parts of the experiments, behavioral confirmation and selective incentives were introduced as monetary side-payments. In this way we tried to model real situations, in which related individuals experience direct social control. As we deal with unrelated subjects in the experiments in artificial circumstances, it was only possible by transforming the utility of these social incentives into monetary rewards. Obviously, we predict these effects to be stronger. With their introduction, however, our main intention was to provide a meaningful comparison between the size of monetary and internalized social incentives.

A deeper question concerns why do people experience social control and why do they experience social incentives in relation to strangers. Evolutionary psychology offers an answer in the ancestral environment (Cosmides and Tooby, 1992). At a certain time of human evolution, survival success and adaptation to the environment was increased by the emergence of human sociality. There is a growing literature that helps to explain the evolution of early human sociability by using computer simulation or game theoretic tools (e.g., Axelrod, 1986; Allison, 1992; Young, 1993; Binmore, 1998; Bowles and Gintis, 1998; Bateson, 2000). Mainstream sociology, however, concentrates not on natural selection, but on culture and norms that are learnt as “rules of the game” during the socialization process. It is not our goal to be involved in this debate on nature and nurture (for a synthesis, see Gintis, 2001). We are not interested in the origins of social incentives; we just aim to test whether they operate also in an impersonal environment.

From previous experiments we know that many other factors influence cooperation rates in social dilemmas (e.g., Ledyard, 1995). These factors might be relevant also in intergroup related collective action. With the introduction of personal characteristics as control variables into the analysis, we will be able to test the significance of “traditional” research variables as gender, study direction, experience with similar experiments, risk preferences, and social orientations. We show whether prosocial individuals are more concerned about harmful outcomes and thus abstain from contribution or whether they have higher contribution propensities and are even the initiators of harmful intergroup conflict.

We have to emphasize that decision algorithms that we discussed in Section 2.4 (dominant strategy, dominant reply, local common knowledge, and expected value rules) are difficult to be traced in the experiments. Looking only at the monetary rewards, subjects have no dominant strategies in the payoff structure of the experiment. We are not able to test the hypothesis that some subjects might evaluate social control so much that contribution becomes unconditionally their best choice. Furthermore, subjects have no information about the utility function of neighboring others as it is assumed partly in the dominant reply, the local common knowledge, and in the expected value rules. Similar problems hold for subjective expected utility rules (cf. Esser, 1986) and other belief-based models (Camerer and Ho, 1999).

In this section we provided an introduction to the experimental analysis of intergroup competitions. We discussed motivations to investigate structural effects that until now have been neglected in the experimental literature. In Section 3.2 we formulate our hypotheses for the experiments. In Section 3.3 we introduce the new experimental design. The model we use to test our hypotheses is specified in Section 3.4.1. A discussion about which control variables are included in the analysis can be found in Section 3.4.2. Section 3.5 describes the method with the discussion of some major characteristics of the subjects, the experimental procedure, and some technical difficulties. Results in Section 3.6 contain descriptive statistics and the discussion of questionnaire data, followed by the multilevel analysis of contribution propensities. Conclusions are drawn in Section 3.7.

3.2 Research questions and hypotheses

The experiments were designed to answer research questions about the effect of structural embeddedness on intergroup conflict. At the aggregated level we aim to answer whether harmful intergroup outcomes more likely in sessions in which members of the two groups are seated in a segregated pattern. With regard to the underlying micro mechanisms, our main research question concerns whether social control affects individual decisions in a direct and in an internalized form. These questions will be answered by analyzing individual decisions in a series of single-shot IPG games.

The general conclusion of previous IPG experiments is that individuals often make sacrifices for their group. In our experiments we will test whether higher contribution rates can be partly explained by effects of social control. Furthermore, we are interested whether eye contact between the subjects is sufficient to trigger internalized forms of social control in impersonal experimental conditions. We will test the presence of three types of social control mechanisms: selective incentives (s), behavioral confirmation (b), and traitor rewards (t). When internalized, these forms of social control create non-monetary incentives for the subjects, which can be expressed as part of their utility functions. Since both monetary and these non-monetary incentives enter the game, the “real” payoff matrix is not equivalent to the “monetary” payoff matrix. The “real” game could have multiple equilibria (cf. Gächter and Fehr, 1999). Contribution, for instance, could be a dominant strategy (cf. inequality 2.4.1). Formally, in the experimental game contribution is a dominant strategy of individual *i,* if assuming a linear utility function on rewards and linear effects of network ties,

$$f_i(s_i - b_i) > g_i t_i + e_i \tag{3.2.1}$$

holds, where e_i stands for the reward for free riding, f_i denotes the number of fellow and g_i the number of opposite neighbors of i. The relative weight of the utility of monetary rewards and of the utilities attached to different forms of non-monetary incentives can change from person to person.[1] We do not assume a specific form of utility function that can be applied to everyone. However, we are interested only in mean tendencies and in trying to estimate the average relative size of different forms of social incentives.

Our main interest is in the analysis of social control that is experienced in dyadic contact. We predict that social control enters subjects’ considerations in an internalized form when eye contact is established and in a direct form when monetary side-payments are included. With regard to different forms of social control, we can explicate the following hypotheses. These hypotheses are derived directly from the discussion in Chapter 2.

Selective incentives always reward contribution or punish defection (Olson, 1965).

<u>Selective incentives</u>: *Selective incentives both in an internalized and in a monetary form have a positive effect on contribution propensities.* More connections to group fellows means the distribution of selective incentives from multiple sources. *Hence, the higher the number of group fellows connected, the higher the contribution rate is.*

On the other hand, the effect of behavioral confirmation is not only dependent on the composition of the network neighborhood, but also on expected decisions of neighbors. We presume that subjects do not make qualitative differences between fellow neighbors. We can therefore derive the following hypothesis:

[1] For this reason, we use the subscript *i* for non-monetary incentives in inequality (3.2.1). Later on, this subscript will be omitted.

Behavioral confirmation: *We predict that behavioral confirmation both in an internalized and in a monetary form have an effect on contribution propensities. The direction and the size of the effect depend on the number of expected contributors and on the number of expected defectors among group fellows connected. If the former is higher, the effect is positive. If the latter is higher, the effect is negative. The higher the difference, the higher the size of the effect is.*

The presence of neighbors from the opposite team triggers an internalized form of social control we called traitor rewards. Because of research limitations and for the sake of simplicity, we did not introduce these incentives in a monetary form in the experiments.

Traitor rewards: *We predict that internalized traitor rewards have a negative effect on contribution propensities. The higher the number of members of the opposite group connected, the lower the contribution rate is.*

Network connections were conceptualized as adjacency in the seating configuration in the experiment. As we believe that neighbors are the direct source of social control, different neighborhood compositions would lead to different contribution propensities. At the aggregated level, higher contribution propensities mean higher likelihood of intergroup conflict. As a consequence, different outcomes can be predicted for different neighborhood structures. Theoretical analysis in Chapter 2 has shown that from the nature of the specified social control mechanisms it follows that segregation is likely to promote intergroup conflict. On the basis of this result, we formulate the following hypothesis about the macro consequences of dyadic social control in the IPG experiments:

SEGREGATION HYPOTHESIS: *In a segregated structure, contribution rates will be higher and intergroup conflict will be more likely.*

Furthermore, the analysis in Chapter 2 also specified the impact of the relative size of social control mechanisms on intergroup conflict. Results demonstrated that the segregation effect on intergroup conflict is stronger where selective incentives are relatively important when compared to behavioral confirmation. In order to test this theoretical prediction, a *normative pressure* condition and a *confirmation pressure* condition was implemented in the experiments. In the normative pressure condition, selective incentives were introduced as monetary side-payments. In the confirmation pressure condition, monetary behavioral confirmation rewards were included. On the basis of the theoretical prediction, the hypothesis we formulate for the experiments is as follows:

The segregation effect on the likelihood of intergroup conflict will be stronger in the normative pressure condition than in the confirmation pressure condition.

In this section, we summarized our main research questions and hypotheses for the experiments. In the next section, we introduce a new experimental design that is invented to test these hypotheses.

3.3 Experimental design

In this section, the structure of the extended IPG game used in the experiments will be outlined. There were two groups: the red group and the green group consisting of five members each. Every player decided individually whether to keep a bonus of 11 points completely (in the experiments, 1 point was equivalent to 1 NLG = 0.42 USD) or to give all of it to help their group in the competition. Depending on the number of contributors in the groups, public good and “bad” rewards were distributed equally among all group members. The sizes of these rewards in the experiments are shown in Table 3.3.1.

Everyone in the groups received these rewards, regardless of the decision to keep or give away the bonus of 11 points. Table 3.3.1 does not include the bonus reward that is added to the payoff of those subjects who decided to keep it. Moreover, to ensure positive payoffs, every subject was entitled to 15 points additional payment at the end of the experiment.

Table 3.3.1 The IPG game (with clash punishment) used in the experiments

payoffs 1 point = 1 NLG		*number of contributors in the green group* 0	1	2	3	4	5
number of contributors in the red group	0	0 / 0	0 / 0	0 / 0	15 / -15	15 / -15	15 / -15
	1	0 / 0	0 / 0	0 / 0	15 / -15	15 / -15	15 / -15
	2	0 / 0	0 / 0	0 / 0	15 / -15	15 / -15	15 / -15
	3	-15 / 15	-15 / 15	-15 / 15	-11 / -11	15 / -15	15 / -15
	4	-15 / 15	-15 / 15	-15 / 15	-15 / 15	-11 / -11	15 / -15
	5	-15 / 15	-15 / 15	-15 / 15	-15 / 15	-15 / 15	-11 / -11

Note: The payoffs are public good rewards distributed to everyone in the red (bottom left corner of each cell) and in the green (top right corner) group.

Due to Definition 1.2.1.1, in intergroup conflict the collective interest of a group is hurt by the collective action of the other group. In the experiment, if there were at least three contributors in one or both of the groups, negative rewards were distributed (cf.

Table 3.3.1). These are the combinations (victory of one side or clash) we call *intergroup conflict* in the experiment.

There are three pure strategy Nash equilibria in this game. One is when nobody contributes. In this situation, no player can be better off by giving the bonus away since the outcome cannot be changed by a single contribution. The other equilibria are when there are three contributors in one group and none in the other team. In this situation, contributors do not have an incentive to switch because the reward for victory outweighs the value of the bonus. Free riders are happy in the winning group and alone they cannot change the situation in the losing group. Since communication is not allowed between the players, there are strong arguments against the occurrence of the latter equilibria. First, nobody would be really interested in making the sacrifice when he or she could also free ride on the effort of other team members. Second, making such a sacrifice does not mean success immediately; there should be at least two other fellows who are willing to make this decision. This makes contribution a very risky alternative and makes the latter equilibria vulnerable (cf. Selten, 1975). On the other hand, overall abstaining from contribution is a stable equilibrium. Although keeping the bonus is not a dominant strategy in the game, it nevertheless provides, almost under every condition, a better individual payoff.[2]

In order to obtain more data in the experiments, the game was played many times in each session, but subjects received payments in a randomly selected single round only. No information was provided during the experiment about what has happened in earlier rounds. This method was applied in earlier team game experiments by Bornstein and Ben-Yossef (1994).

To have a low variation in neighborhood composition and to facilitate the testing of the segregation hypothesis, only simple neighborhood structures were applied in the experiments. Every experiment started with a control condition, in which subjects had to make their decisions in isolation and without the knowledge of their group membership. After the control condition, color labels were introduced and subjects were seated due to three structural patterns that are shown on Figure 3.3.1. These structural conditions varied in between the experiments. In all structures, subjects had at most two neighbors. The structural position of the groups was completely symmetrical in each case.

Seating patterns were varied between experimental sessions. Visibility conditions were varied and direct social incentives were introduced within every experiment. The latter was done in order to map neighborhood effects and to provide a meaningful comparison for the relative size of the effect of internalized social incentives. The importance of internalized social incentives varies across individuals. We are only

[2] The exceptions are the following situations. A member of the red group is better off with contribution, if without this player there are {2;0}, {2;1}, {2;2}, {3;3}, or {4;4} contributors in the red and in the green group, respectively.

able to demonstrate the relative importance of internalized social control on average and the extent of variation between subjects.

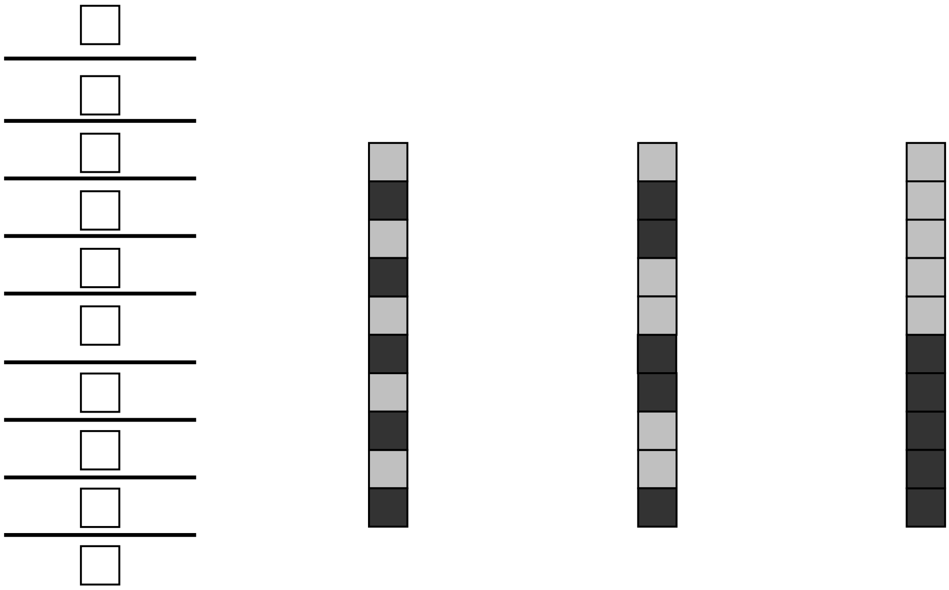

Figure 3.3.1 Structural conditions in the experiment: control condition, low, medium, and high clustering

Notes: Light and dark colors indicate seats of the members of the red and green group. In the control condition, no color labels were introduced.

Social control of the network neighborhood is predicted to be strongest in the presence of direct monetary rewards. Between experimental sessions, two conditions were applied. Next to the payoffs that were present in the beginning of the experiments (see Table 3.3.1), in the *confirmation pressure* condition direct behavioral confirmation incentives (b=5 NLG), in the *normative pressure* condition direct selective incentives (s=5 NLG) were introduced at a certain point in the experiment. Later, in both conditions the other type of incentives was also introduced. Subjects received 5 NLG behavioral confirmation reward if one of their fellow neighbors chose the same action as they did and received 10 NLG if two of their fellow neighbors acted on the same way. Selective incentives were distributed regardless of the decision of neighbors. Contributing subjects received 5 NLG for each fellow neighbor they had. In the low clustering condition there was no change due to the absence of fellow neighbors. This way we had a 2×3 block-design of experimental conditions (see Table 3.3.2).

Table 3.3.2 The number of sessions by experimental conditions

level of clustering	**low**	**medium**	**high**
confirmation pressure (*b first*)	*3*	*4*	*4*
normative pressure (*s first*)	*3*	*4*	*4*

Every experiment consisted of four parts (see Table 3.3.3). All experimental sessions started with an anonymous control condition (Part I) in which subjects were isolated from each other. They did not see each other and did not know which group they belong to. This part allows us to estimate the baseline contribution rates without eye contact between the subjects.

Table 3.3.3 Overview of experimental parts

Part I	anonymous control condition
Part II	eye contact is established
Part III	one form (*b*/*s*) of social control is introduced in a monetary form
Part IV	the other form (*s*/*b*) of social control is introduced in a monetary form

3.4 The model for explaining contribution propensities

3.4.1 Main effects: internalized social control

In this section, we describe the model we use to test our hypotheses in the IPG game experiments. In the explanatory model of individual contribution propensities we incorporate different forms of social control that are believed to be the underlying mechanisms of the segregation effect on intergroup conflict at the macro level. Besides social control being the focus of our interest, we briefly discuss the influence of personal characteristics that are handled as control variables.

For the analysis of experimental data we will use multilevel logistic regression (Bryk and Raudenbush, 1992; Goldstein, 1995). In our two-level model, single decisions are the lower level observations and characteristics of the subjects, who took these decisions, are the group level observations. The two-level model allows us to correct for the methodological problem that observations within the subjects are not independent. For the binary dependent variable, we will use the logit transformation. Formally, let the function P_{ri} denote the propensity of actor i to cooperate in the rth single-shot game. The propensity of cooperation is specified by the logit link function (Goldstein 1995: Chapter 7), which is the natural logarithm of the quotient of the probability of contribution $P_{ri}(C)$ and the probability of defection $P_{ri}(D)$.

$$P_{ri}^{I} = \ln\left(\frac{P_{ri}(C)}{P_{ri}(D)}\right) = \alpha_0 + \varepsilon_i + \xi_{ri}, \qquad (3.4.1.1)$$

where α_0 is the baseline contribution propensity. From previous experiments we know that α_0 depends primarily on the payoff parameters, on the criticalness of the decision (which is a function of group size and the threshold of collective action) and on experimental conditions of confidentiality and anonymity (whether subjects know each other, whether they are in the same room, etc.). Notation ε_i stands for a subject level error term and ξ_{ri} is intra-individual variation. The latter term represents the residual variance that is not estimated in models that include the random intercept α_0. We assume that the subject level error has a zero expected value and has a normal distribution, that is

$$\varepsilon_i \sim N(0, \sigma^2),$$

where the variance σ^2 is going to be estimated. This baseline model does not contain any explanatory variables and allows us to model behavior in the anonymous control condition (Part I).

The interindividual variation of contribution propensities depends on personal characteristics, like gender, study direction, experience in similar experiments, attitudes towards risk, or social orientations. We will include personal characteristics in the analysis as control variables. They enrich research with interesting insight and comparisons can be made with previous findings. They might also interact with internalization of social control. These effects will be discussed in the next section.

Intra-individual variation results from experimental manipulations and also because subjects might reconsider their expectations about the behavior of others. These main factors are relevant after the introduction of minimal contact in Part II. Additional reasons for such variation that can already be present in the control condition are stochastic individual decisions, consideration of mixed strategies, or simply inconsistency. In our simplest model, we assume that intra-individual variation is not correlated with the round number r and has a zero expected value. However, we have to modify this assumption and add a trend element if we find indications of learning the structure of the game through the experiment.

With the introduction of minimal contact (Part II), we predict that internalized social incentives affect individual decisions. The propensity of cooperation will also be dependent on internalized selective incentives, behavioral confirmation, and traitor rewards, as expressed by equation (3.4.1.2):

$$P_{ri}^{II} = \alpha_0 + s_0 f_i + b_0\left(\hat{f}_{cri} - \hat{f}_{dri}\right) - \left(t^+ + t^-\right)g_i + \varepsilon_i + \xi_{ri}. \qquad (3.4.1.2)$$

To recall, notation f_i stands for the number of fellow neighbors of i and g_i for the number of neighbors from the opposite group. Parameter s_0 denotes internalized selective incentives (prestige, respect, etc.) and b_0 stands for internalized behavioral confirmation rewards. We need to estimate these parameters from the experimental results. The expression after b_0 within the brackets denotes the difference between the expected number of contributing and defecting fellow neighbors in decision round r. If the latter is higher, behavioral confirmation decreases the likelihood of contribution. We used a direct way to estimate the expectations of subjects about the number of contributing and non-contributing neighbors. Subjects were asked to forecast the decision of their left and right neighbors before every decision round.

The number of neighbors from the opposite group affects contribution rates through traitor rewards. These are positive (t^+) and negative (t^-) selective incentives rewarding defection and punishing contribution.[3] For the sake of simplicity, we denote their sum by t_0. In the simplest model, we only estimate the average individual importance of internalized social control, therefore the subscript i is omitted for estimates s_0, b_0, and t_0. However, in part of the analysis, we will consider that there is a random variance in the size of these effects. Specifically, we will assume that the effects of internalized social control for the subjects are normally distributed around their means. This is consistent with the statement that individuals do not assign the same relative utility for social control, but the utilities are scattered normally around a certain mean evaluation. In this part of the analysis, we will estimate the variances of the influences of different forms of internalized social control, as well as their covariances.

Finally, when explicit monetary payoffs of social control (s and b) are introduced (Parts III and IV), we have to extend equation (3.4.1.2) with additional terms. The simple model in its general form can be written as

$$P_{ri} = \alpha_0 + \left(s_0 f_i + b_0\left(\hat{f}_{cri} - \hat{f}_{dri}\right) - t_0 g_i\right)p^{II} + s_1 f_i p^s + b_1\left(\hat{f}_{cri} - \hat{f}_{dri}\right)p^b + \varepsilon_i + \xi_{ri}. \tag{3.4.1.3}$$

The p^j dummies denote experimental parts: p^{II} indicates whether or not minimal contact is present, p^s denotes whether or not direct selective incentives are introduced, and p^b shows the presence of direct confirmation rewards. We have to emphasize that s_1 and b_1 in equation (3.4.1.3) are parameters we need to estimate and therefore we distinguish them from the monetary values s and b. The utility of these monetary rewards might differ between subjects, therefore, in part of the multilevel analysis we will assume a normal distribution of the estimates over the subjects.

3.4.2 Control variables and interaction effects

Instead of including individual effects as dummies, it is more meaningful to analyze what kind of personal characteristics have an influence on contribution propensities. We do not derive hypotheses about the effects of these variables, but we include them as controls because they might have a strong impact on behavior in the experiments.

[3] In Chapter 2, we considered positive traitor incentives only. The consideration of both positive and negative traitor rewards does not add complications to the model.

We know from previous social dilemma experiments that gender is an important predictor of cooperation. However, findings are contradictory about whether women or men are more cooperative (e.g., Isaac, McCue, and Plott, 1985; Mason, Phillips, and Redington, 1991; Frank, Gilovich, and Regan, 1993; Brown-Kruse and Hummels, 1993; Nowell and Tinkler, 1994; Cadsby and Maynes, 1998; Ortmann and Tichy, 1999). Most subjects participating in experiments are students at different faculties of the university. Direction of study might cause individual differences in willingness of contribution. Previous research found that economists have lower contribution rates (Marwell and Ames, 1981; Carter and Irons, 1991; Frank, Gilovich, and Regan, 1993), although there are also experiments that do not find this effect (Isaac, McCue, and Plott, 1985; for an overview, see Ledyard, 1995: 161, 179).

Besides these background variables, we control for attitude measures that indicate special forms of individual utility functions. Previous findings show that attitudes towards risk correlate with contribution propensities (Suleiman and Or-Chen, 1999). Since the contribution decision involves the possibility of a higher reward, but also involves the risk of losing the bonus completely, subjects with a risk-seeking attitude might have higher contribution rates (Budescu, Rapoport, and Suleiman, 1990). On the other hand, there are arguments that in repeated social dilemmas risk aversion increases cooperation (Raub and Snijders, 1997; van Assen and Snijders, 2002). In our experiments, we use attitudes towards risk only as control variables. For the measurement of risk preferences, questions with preference comparisons (see Farquhar, 1984) were used (see Appendix 3.A.2).

Utility functions can also include altruistic elements, which certainly influences rational decision making in social dilemma experiments (e.g., Liebrand, 1984; Doi, 1994). Subjects, who order positive utilities for the gains of others, behave differently from individualistic ones. For the approximation of such utilities, standard questions regarding social orientations were used. They consisted of a series of decomposed games with an unknown person.[4] We presumed that individuals are only *prosocial* (cooperative), *individualistic*, or *competitive*. Previous research found only these types relevant in describing human behavior (van Lange et al., 1997; van Lange, 1999; Suleiman and Or-Chen, 1999). Among each type we distinguished an egalitarian tendency (cf. van Lange, 1999). Although in a two-person PD game or in a public good experiment we should predict higher contribution rates from prosocial subjects, it is not so evident in the IPG game. One could argue that subjects who order utility weights for rewards of unknown others, would do this equally for everyone, including out-group members. Consequently, their contribution rates would not be different from individualistic subjects. A counter-argument is that prosocial (and also egalitarian) orientation is associated with high utility for *social identity*, which is obtainable in a relational comparison with the out-group. Hence prosocial orientation is primarily directed towards in-group members.

[4] The exact questions can be found in Appendix 3.A.1.

In part of the analysis, we also included interaction effects of background variables and social control, because the relative size of internalized social control in the utility function might depend on certain personal characteristics. There are contradictory findings in previous experiments about whether people are more likely to think of others of the same sex to be contributors and in general, whether men or women are more likely to be thought of as better contributors (Ortmann and Tichy, 1999; Solnick and Schweitzer, 1999). For explorative reasons, we control for interactions between gender and social control. We also include as controls interactions between social orientations and the relative weight of social control.

Experiments were combined with repeated IPG games. Repeated games followed single-shot games in all four experimental parts. We designed the experiments so as to exclude possible influences of previous decisions. Subjects were explicitly told before every part that previous parts and repeated games are completely independent from the next part. We predicted no history effects on single-shot decisions, but as a test of this hypothesis, we included previous outcomes of repeated games as control variables in part of the analysis.[5]

3.5 Method

3.5.1 Subjects

203 subjects took part in the experiments at the University of Groningen in May 2000. Subjects were recruited via e-mail and board advertisements promising monetary rewards for participation in a "team competition" experiment. All 203 subjects completed the decision tasks and only two failed to complete the post-decision questionnaire. Altogether, 21 sessions took place and subjects made 4060 decisions (20 each).

114 (56.2%) subjects were female. 187 (92.1%) subjects were university students at the time of the experiments and 16 had already graduated. Students came from all faculties of the university: 55 studied behavioral or social sciences, 47 subscribed for literary studies or art, 26 studied natural sciences, 17 studied law, 13 studied economics, 10 were students at the business faculty, there were 8 students of medical science, 8 subjects studied spatial sciences, and one subject read philosophy. Because of similarities and for the sake of simplicity, economic, business, and spatial sciences were merged in the analysis (furthermore, these faculties have the same physical location) and the student of philosophy was allocated to the category of literary studies and art. The study direction of two subjects was unknown.

Single-shot games (only the decision rounds) took approximately three minutes in each experimental part. During this time subjects had to make five decisions. Counting

[5] In the repeated games, subjects were informed about the result of the previous round.

all four parts of the experiment, every subject played in 20 single-shot games. The entire experiment was on average 80 minutes long.

The payoff for subjects was contingent on their decisions, as well as on the decisions of other participants of the session. Individual payoffs were calculated on the basis of outcomes in the single-shot and in the repeated games. The weighted average of the payoffs obtained in the decision rounds was paid to the subjects after the experiment in sealed envelopes. Payoffs from every round of the repeated games had a single weight. From the single-shot games, only one was selected randomly in each experimental part to be included in the calculation. This payoff had a weight of five rounds (the number of single-shot games in one experimental part). Payoffs varied between 14 and 32 Dutch guilders with an average of 21.1 NLG. Reserve subjects received 10 NLG for their appearance. If subjects ran out of decision time, a random decision was implemented. For all such cases, final payment was decreased by 1%. Out of 4060 decisions this happened only 26 times (0.64%). Random decisions are not included in the analysis.

3.5.2 Procedure

Experiments were conducted in the same computer laboratory.[6] Upon arrival, subjects were randomly seated at a computer. Removable walls separated the subjects to ensure their privacy. Subjects received instructions on paper and on their screen. After reading the instructions they were allowed to ask the experimenter questions. After the questions had been answered, subjects were not allowed to talk. All participants strictly adhered to the rules. After the questions, an examination of understanding followed.

Subjects understood the task quite well, on average they answered 16.5 questions in five minutes, from which 13.7 (83%) were right. The mean proportion of correct answers was 80% with a standard deviation of 18%. Only nine subjects gave more wrong than right answers, one of the nine gave answers that could be considered random. One subject did not answer any of the quiz questions.

Every experiment consisted of four parts. In each part, subjects played five rounds of single-shot IPG games, followed by a number of repeated games. Experiments began with an anonymous start (Part I). Group membership was not yet announced. In every decision round, subjects had to decide whether they would keep the 11 NLG bonus or give it to help their group to achieve success in the competition. These two options appeared in a randomized order on their screen. The bonus was represented also graphically as a bag of money. Subjects were assured of the anonymity of their decisions and that they would receive any money they earned during the experiment in sealed envelopes, after the experiments had ended. In the single-shot games, it was announced that every decision counts towards the final payment, but that only one game of each part would be chosen randomly for payment.

[6] The computer program for the experiment was written by Sicco Strampel in Delphi 5.

In the beginning of Part II, separator walls were removed and group membership was made public by the experimenter. Red and green flags were attached to the monitors and subjects also received an A-4 colored paper with the color of their group. In each condition, subjects were seated behind computers due to the neighborhood configuration of the given session. Subjects were explicitly asked to look around in the laboratory and then played five rounds of the same IPG game again. Before every decision in Part II, III, and IV, subjects had to give their expectations about the subsequent decision of their neighbors. The five single-shot games were followed by repeated games.

In Part III, monetary payoffs for social control were introduced explicitly. In 10 sessions (see Table 3.3.2) monetary confirmation rewards and in 11 sessions monetary selective incentives were incorporated. In Part IV, the other type of social control was also introduced in an explicit monetary form (see Table 3.3.3). As in the low clustering condition there were no fellow neighbors, this condition was used to control for independent learning effects (there was no change between Part II, III, and IV).

Calculation and announcement of the individual results followed the experiment. Meanwhile subjects were asked to fill in a questionnaire on their computer. Monetary payments were supplied in sealed envelopes. The first subject, who had completed the questionnaire, could go immediately to the experimenter to receive payment. Other subjects had to wait until they got a signal from the server. Hence, subjects left the laboratory individually, with a short time difference between their departure. They were informed about the aim of the study after the experiments.

3.5.3 Difficulties

The intended number of participants was ten in all the 21 experimental sessions. On average, thirteen subjects were invited to the sessions as it was anticipated that some would not come. Four sessions failed to be completely filled. In these cases, computer players were included in the teams. Subjects were told that they are programmed in a way to resemble human behavior. In fact, they were simple programs playing mixed strategies with condition-dependent probabilities of contribution. Human decisions in the incomplete experiments are also included in the analysis, but computer decisions are excluded. The inclusion of simulated participants did not have a significant influence on the behavior of subjects in the IPG games.[7]

Additional complications occurred due to a computer failure in three experimental sessions. In these cases, the given subject was seated further away from the others at another computer. No neighbors were defined for these subjects. These complications did not cause distortions in the data.

[7] This variable is not significant when added to any of the multivariate models discussed in Sections 3.6 and 5.6.

Some of the participants knew each other. As acquaintances might influence actual decisions in the experiment, we included the number of acquaintances in the experiment as a control variable in the analysis. Not only the number, but also the exact seating location of acquaintances can be a relevant factor. However, inclusion of such variables would add substantial complexity to the explanatory model without much theoretically grounded gains.

3.6 Results

3.6.1 Contribution rates under different experimental conditions

Let us first provide an overview of the general results of the experiments. Altogether 420 single-shot games were played (20 decision rounds in each of the 21 sessions). There was no tendency towards equilibria as only one outcome was {0;0}, three were {3;0}, and five were {0;3}. In the control condition, in which no minimal contact was introduced and group membership was not announced, there were no {0;0} or {3;0} outcomes and {0;3} occurred only twice (out of 105 outcomes).

Table 3.6.1.1 Outcomes by clustering conditions in all single-shot games

clustering condition in the experiment	*outcome of the decision round* peace	conflict	Total
control condition	26.97% (271)	73.03% (734)	100% (1005)
low clustering	50.23% (428)	49.77% (424)	100% (852)
medium clustering	13.75% (160)	86.25% (1004)	100% (1164)
high clustering	11.85% (120)	88.15% (893)	100% (1013)
Total N	24.27% (979)	75.73% (3055)	100% (4034)

Notes: Cases in parentheses are weighted (multiplied) by the number of human decisions in the given game. For the χ^2 –test unweighted outcomes are used, N=420.

As a consequence of dyadic social control, we predicted different outcomes by clustering conditions. The segregation hypothesis predicted that conflict will be least likely in the low clustering condition and will be most likely in the high clustering condition. Table 3.6.1.1 summarizes the experimental outcomes by clustering conditions. The hypothesis that the outcomes of the IPG game are independent of clustering conditions can be rejected ($\chi^2(3)=46.370$, $p<0.001$).

Table 3.6.1.1 shows that conflict was already quite likely the outcome in the control condition. It indicates that many subjects had a relatively high baseline contribution propensity. Conflict was much less likely in the low clustering condition, and occurred most often in the high clustering condition, which supports the segregation hypothesis. On the other hand, conflict was almost as likely in the medium clustering condition as in high clustering. Conflict occurred in 85.83% of the cases in the medium and 88.57%

of the cases in the high clustering condition (from unweighted outcomes; t=0.613, two-tailed p=0.541), which is counter to the segregation hypothesis.

Contribution rates by clustering conditions are summarized in Table 3.6.1.2. The hypothesis that contribution rates are the same across different clustering conditions can be rejected (ANOVA $F(3, 4030)$=52.629, p<0.001). This difference is the result of internalized *and* direct social control. In order to test whether internalized social control can alone cause such differences between clustering conditions, we compare the results from Parts I and II. The comparison reveals that eye contact made an increase in contribution rates. The difference is significant at the 5% level, but not at the 1% level (t=1.722, one-tailed p=0.043). However, in Part II, the contribution rate was highest in the medium clustering condition, which contradicts the segregation hypothesis. The hypothesis that contribution rates are the same in Part II cannot be rejected (ANOVA $F(3, 2003)$=1.005, p=0.366, not significant). Since there is an increase in contribution propensities in the medium and high clustering conditions compared to the control condition, social control might exist also between fellow neighbors in an internalized form. In order to get a better indication of whether or not internalized social control is activated when eye contact is established between subjects, we will use a multivariate analysis and estimate parameter values of the model expressed in equation (3.4.1.3).

Table 3.6.1.2 Average contribution rates (%) in different clustering conditions and parts of the experiment

	incentives introduced first	**low clustering**	**medium clustering**	**high clustering**	**Total**
*Part I**		49.64 (280)	51.81 (386)	46.61 (339)	49.45 (1005)
Part II		50.35 (282)	55.84 (385)	52.84 (335)	53.29 (1002)
P III	*b* (*confirmation pr.*)	-	58.42 (190)	47.33 (150)	53.53 (340)
	s (*normative press.*)	-	63.82 (199)	75.66 (189)	69.59 (388)
	Part III total	40.35 (285)	61.18 (389)	63.13 (339)	55.97 (1013)
P IV	*b* (*confirmation pr.*)	-	62.63 (190)	68.00 (150)	65.00 (340)
	s (*normative press.*)	-	71.00 (200)	81.48 (189)	76.09 (389)
	Part IV total	25.96 (285)	66.92 (390)	75.52 (339)	58.28 (1014)
Total (without Part I)		38.85 (852)	61.34 (1164)	63.87 (1013)	55.86 (3029)
Total		41.52 (1132)	58.97 (1550)	59.54 (1352)	54.26 (4034)

Notes: The number of cell-relevant cases is in parentheses. All human decisions are included.

* In Part I, subjects did not know their group membership and they did not see each other. Therefore their assignment to the different clustering conditions only illustrates baseline contribution rates in the different experimental sessions.

Table 3.6.1.2 also shows average contribution rates in Parts III and IV of the experiment. The hypothesis that contribution rates are the same in the different parts can be rejected (ANOVA $F(3, 4030)$=5.869, p=0.001). We also tested whether the introduction of different direct social incentives in Part III and IV made a difference in contribution rates or not. The hypothesis that contribution rates are the same in the different conditions can be rejected both in Part III (ANOVA $F(2, 1010)=30.800$, $p<0.001$) and in Part IV (ANOVA $F(2, 1011)=108.721$, $p<0.001$). Our hypothesis in Section 3.2 predicted that the introduction of monetary selective incentives would result in higher contribution rates than when behavioral confirmation is introduced in Part III ($t=4.487$, one-tailed $p<0.001$). *Results confirm this hypothesis*. Furthermore, earlier introduction of normative pressure made a difference also in Part IV ($t=3.285$, two-tailed $p=0.001$). This result indicates that history effects still play a role in determining individual decision, despite the lack of feedback regarding the results of single-shot games. We will return to this point at the multilevel analysis of results. Furthermore, figures in Table 3.6.1.2 also *support the hypothesis that under normative pressure the effect of segregation is stronger than under confirmation pressure*. In Part III, under normative pressure average contribution rates are higher in the high clustering condition (75.66%) than in medium clustering (63.82%). On the other hand, under confirmation pressure average contribution rates are higher in the medium clustering condition (58.42% vs. 47.33%).

3.6.2 Questionnaire data

Questionnaire data can only provide partial support for the results. Most subjects reported that they felt important differences between experimental parts, but we are not able to establish the mechanisms behind these differences from the answers. The analysis of actual behavior more convincingly shows the important factors that affect individual decisions.

Table 3.6.2.1 reports some descriptive statistics of relevant questions. The question on the importance of minimal contact (first column) was formulated as: "In the beginning of the experiment walls were separating you and your neighbor(s). Later the walls were removed and flags were attached to the monitors to indicate the group membership of participants. When you had to make your decisions, how important was this difference for you?" Most subjects answered that it was a rather important difference, but there were also many participants for whom this was not important.

Other questions reported in Table 3.6.2.1 were formulated similarly, but they were related to the difference between Part II and Part III and to the difference between Part III and Part IV, respectively. We display descriptive statistics about the subjective importance of these differences separately for sessions, in which direct selective incentives were introduced first (second and fourth columns) and for those, in which direct behavioral confirmation was introduced first (third and fifth columns). Only answers of those subjects who had a fellow neighbor are reported as only they received these new incentives. The hypothesis that subjects felt one form of social control more important at the first introduction can be rejected (Wilcoxon rank test $W=3721.5$, two-tailed $p=0.135$). Furthermore, the difference between these sessions is not significant with regard to changes made before Part IV (Wilcoxon $W=3918.5$, $p=0.598$). This is a somewhat surprising result, since contribution rates differed significantly between these sessions both in Part III and in Part IV (cf. Table 3.6.1.2), which implies that subjects' opinions often do not match their behavior.

On the other hand, the difference between the mean rank and the distribution of those who had these new incentives and those who had no new incentives is highly significant (Wilcoxon W=9332.5, p<0.001; Kolmogorov-Smirnov Z=5.023, p<0.001 for the first introduction and Wilcoxon W=9151.0, p<0.001; Kolmogorov-Smirnov Z=5.126, p<0.001 for the second introduction).

Table 3.6.2.1 Frequencies of answer categories of questionnaire data

questions	*importance of eye contact*	*importance of selective incentives in Part III*	*importance of behavioral confirmation in Part III*	*importance of behavioral confirmation in Part IV*	*importance of selective incentives in Part IV*
very import. (1)	28	13	15	22	22
important (2)	67	30	30	30	25
neutral (3)	42	18	12	12	9
not important (4)	40	8	2	5	3
not imp. at all (5)	24	1	2	1	2
total answers	201	70	61	70	61
mean	2.83	2.34	2.11	2.04	1.98
st. deviation	1.24	0.96	0.93	0.95	1.01
median	3	2	2	2	2
mode	2	2	2	2	2

Other questions were related to the main intentions of subjects during decision making in the experiment. These answers, however, were very much influenced by the repeated games; therefore it would not be adequate to report them here.

3.6.3 Analysis of contribution propensities: a simple model

To understand the underlying mechanisms of the segregation effect on intergroup conflict, we analyzed individual decisions. The first model in Table 3.6.3.1 reports results for the two-level model expressed in equation (3.4.1.3). For the binary dependent variable of individual contribution, the logit transformation was used. The second model assumes that estimates of social control over subjects are normally distributed around their mean. In this model the variances and covariances are estimated as random effects. All human decisions except 23 cases (0.006%) are included. In these 23 cases subjects did not present any expectations about the behavior of their neighbors. In total, 4011 decisions are included in the analysis for 203 subjects.

The two models provided similar estimates. All effects are in the predicted direction. Hypotheses about the existence of internalized behavioral confirmation and internalized traitor rewards are confirmed. The effect of the third type of internalized social control (selective incentives) was not significant. As predicted, both forms of direct social control had a significant effect.

Table 3.6.3.1.a. Results of multilevel logistic regression on contribution propensities

independent variable	hypothesis about the direction of effect	multilevel model with fixed slopes of main effects	multilevel model assuming random slopes of social control effects
FIXED EFFECTS			
α baseline contribution propensity	?	-.038 (.082)	-.037 (.082)
s_0 internalized selective incentives	+	.109 (.072)	.117 (.072)
s direct selective incentives	+	.407*** (.088)	.363*** (.104)
b_0 internalized behavioral confirmation	+	.617*** (.065)	.640*** (.077)
b direct behavioral confirmation	+	.619*** (.104)	.615*** (.118)
t_0 internalized traitor rewards	-	-.175** (.055)	-.173** (.057)
RANDOM EFFECTS			
interindividual variance σ^2		.616^{+++} (.085)	.628^{+++} (.121)
σ^2_{ui} (s_0)			.000 (.000)
σ^2_{ui} (s)			.300^{++} (.139)
σ^2_{ui} (b_0)			.196^{+++} (.093)
σ^2_{ui} (b)			.326^{+++} (.226)
σ^2_{ui} (t_0)			.009 (.050)
Covariances are reported below			+
-2 Log Likelihood model		4480	4430
Improvement χ^2 (df in parentheses)		939*** (5)$^{\#}$	50*** (20)

Table 3.6.3.1.b. Random effects: estimated covariances

σ_{uxy}	ε_i	s_0	s	b_0	b
s_0	.000 (.000)				
s	-.252 (.108)	.000 (.000)			
b_0	.147 (.083)	.000 (.000)	-.194 (.085)		
b	-.359^{++} (.131)	.000 (.000)	.128 (.132)	-.079 (.116)	
t_0	-.005 (.072)	.000 (.000)	.425 (.153)	-.169 (.109)	.176 (.165)

Notes: N=4011 decisions for 203 subjects. Iterative Generalized Least Squares estimates. Numbers in parentheses are standard errors. ** significant at the 1% level, *** significant at the 0.1% level (two-tailed).

For testing random effects it is more appropriate to use deviance tests: $^{++}$ significant at the 1% level, $^{+++}$significant at the 0.1% level (significance of difference in deviance compared to model without random slopes, for random covariates deviance is compared to model without random covariates).

$^{\#}$Basis of comparison: baseline multilevel logistic regression expressed in equation (3.4.1.1); α: 0.174** (0.066); σ^2: 0.674^{+++} (0.087).

Internalized behavioral confirmation had a very strong effect, approximately as strong as 5 NLG monetary incentives for behavioral confirmation. The parameter estimate of internalized behavioral confirmation 0.617 means that an additional fellow neighbor, who is expected to contribute, increases the odds by 85.34%. For instance, in the first model, for an average subject, the predicted likelihood of contribution is 49.05% in the control condition. In the minimal contact condition, the expected likelihood of her contribution, if she were to have one contributing fellow neighbor and no opposite neighbor is 66.55%. For this increase, internalized behavioral confirmation takes the most responsibility.

For testing hypotheses about random effects it is more appropriate to use deviance tests than the *t*-test (cf. van Duijn, van Busschbach, and Snijders, 1999: 192-193). Baseline contribution rates between subjects had a high unexplained variance. The influence of behavioral confirmation and monetary selective incentives varied significantly between subjects. The hypothesis that the sizes of traitor rewards and internalized selective incentives are the same for the subjects cannot be rejected. High positive deviations from the average baseline contribution rate were correlated with negative deviations from the average importance of monetary rewards for confirmation. This is not surprising given that our general prediction was that subjects, who evaluate monetary gains less, would contribute more to the success of their group.

3.6.4 The effect of personal characteristics

To see which personal characteristics are responsible for the high inter-individual variation, in the next part of the analysis we extended the model by controlling for background variables and certain attitude measures. Again, we conducted two analyses: one assuming fixed social control effects without random variation and another assuming a random variation and covariation of these estimates (see Table 3.6.4.1). The inclusion of control variables did not rule out the significant effects of social control variables. The effect of internalized selective incentives remained insignificant.

There was no significant gender effect, although simple descriptive statistics showed that women had higher contribution propensities (55.94%) than men (52.14%). Based also on descriptive statistics, subjects who already graduated were more contributive (61.54%) than students (53.58%). This effect was not significant in the model, as it was ruled out by other variables, mainly by social orientations. The analysis of study direction did not reveal an effect of economics training. In our experiments, students of natural sciences and law had the lowest contribution rates (48.17% and 48.66%, respectively). They were followed by students of economics (51.54%) and medical sciences (53.55%; this was used as a reference category in models of Table 3.6.4.1). Subjects who studied literary studies or art and students of social sciences had the highest contribution rates (56.91% and 55.85%, respectively).

However, the effect of study direction was also ruled out by other variables. The argument that experience matters at all is questioned by the insignificant effect of participating in a similar experiment before. Again, we could probably explain the difference in descriptive statistics (56.14% vs. 51.44%) by selection on attitude measures.

We characterized subjects as strongly risk-averse, if they chose for risk-averse alternatives both in simple and complex gambles. We found that 91 subjects (45.3%) were strongly risk-averse towards gains, 92 (45.8%) were strongly risk-averse towards mixed gambles, and 83 (39.5%) were strongly risk-seeking towards losses. Although we did not try to estimate utilities more precisely, descriptive statistics confirm that most individuals have an S-shape utility function (e.g., Tversky and Kahneman, 1992). Effects of risk-aversion and loss-aversion are not significant in the models.

The only personal characteristics that we measured and found highly significant in explaining contribution propensities were *social orientations*. For questions about social orientations, 77 (37.9%) subjects gave inconsistent answers. Inconsistency was a significant predictor of contribution rates, which is probably related to the relevance of calculation abilities. Among subjects, who gave consistent answers, 76 (61.3%) were prosocial, which is higher than in previous experiments (for an overview see Schulz and May, 1989). As an exception, Liebrand (1984) found a similar high rate in his experiments conducted in Groningen. Our results clearly support the argument that prosocial (and also egalitarian) orientation is primarily directed towards in-group members and therefore increases contribution rates in the IPG game. The strong effects also indicate that social orientations are important predictors of behavior in intergroup situations. Individuals with prosocial and egalitarian attitudes seem to be responsible for the emergence of mutually harmful outcomes.

Table 3.6.4.1.a. Results of multilevel logistic regression on contribution propensities with personal characteristics

independent variable	hypothesis about the direction of effect	multilevel model with fixed slopes of social control	multilevel model with random slopes of social control
FIXED EFFECTS			
Main variables			
α (constant) baseline contribution propensity	?	.568 (.402)	.736 (.389)
s_0 internalized selective incentives	+	.084 (.073)	.112 (.073)
s direct selective incentives	+	.410*** (.089)	.335** (.105)
b_0 internalized behavioral confirmation	+	.628*** (.066)	.644*** (.080)
b direct behavioral confirmation	+	.635*** (.105)	.604*** (.118)
t_0 internalized traitor rewards	-	-.159** (.056)	-.154** (.055)

cont. next page

Personal characteristics and other subject-level variables		
gender (1=male)	-.201 (.139)	-.221 (.134)
student at the university (1=yes)	-.268 (.360)	-.393 (.346)
studies at the law faculty	-.094 (.356)	-.004 (.341)
studies natural sciences	-.016 (.335)	-.030 (.320)
studies economic, business, or spatial science	-.003 (.326)	.095 (.311)
studies social sciences	.118 (.301)	.180 (.287)
student of literary studies or arts	.124 (.308)	.202 (.294)
did a similar experiment before	-.159 (.132)	-.190 (.127)
strong risk aversion towards gains	-.141 (.131)	-.154 (.126)
strong loss aversion	.117 (.130)	.153 (.125)
consistent answers on social orientation questions	-.359* (.176)	-.397* (.169)
prosocial orientation	.479** (.178)	.485** (.171)
egalitarian orientation	.382* (.172)	.386* (.165)
number of acquainted subjects in the experiment	-.089 (.086)	-.095 (.083)
delay (minutes) at the start of the experiment	.011 (.007)	.010 (.007)
quiz questions answered correctly %	-.005 (.004)	-.006 (.004)
RANDOM EFFECTS		
interindividual variance σ^2	$.537^{+++}$ (.078)	$.554^{+++}$ (.115)
σ^2_{ui} (s_0)		.000 (.000)
σ^2_{ui} (s)		$.311^{++}$ (.143)
σ^2_{ui} (b_0)		$.225^{+++}$ (.099)
σ^2_{ui} (b)		$.296^{+++}$ (.221)
σ^2_{ui} (t_0)		.002 (.048)
Covariances are reported below		+
-2 Log Likelihood model	4430	4382
Improvement χ^2 (df) for model in right column		48*** (20)
vs. previous model	50*** (16)	48*** (16)

Table 3.6.4.1.b. Random effects: estimated covariances

σ_{uxy}	ε_i	s_0	s	b_0	b
s_0	.000 (.000)				
s	-.195 (.106)	.000 (.000)			
b_0	.037 (.084)	.000 (.000)	-.213 (.089)		
b	$-.245^{+}$ (.126)	.000 (.000)	.146 (.131)	-.083 (.118)	
t_0	-.035 (.070)	.000 (.000)	.503 (.160)	-.115 (.113)	.163 (.168)

Notes: N=4011 decisions for 203 subjects. Iterative Generalized Least Squares estimates. Numbers in parentheses are standard errors. * significant at the 5% level, ** significant at the 1% level, ***significant at the 0.1% level (two-tailed).

For testing random effects it is more appropriate to use deviance tests: $^{++}$ significant at the 1% level, $^{+++}$ significant at the 0.1% level (significance of difference in deviance compared to model without random slopes, for random covariates deviance is compared to model without random covariates).

There was no significant effect of how many others were acquainted to subjects in the laboratory and delay time at the start of the experiment did not matter for contribution propensities. This result shows that these factors that are related to the experimental environment did not disturb the behavior of subjects.

3.6.5 Learning the structure of the game

Until now we assumed that intra-individual variation (ξ_{ir}) has a zero expected value and it is independent from the decision round r. We based this assumption on the randomized design of single-shot games (Bornstein and Ben-Yossef, 1994). Every decision round could be handled in an equivalent way, since only one of the rounds would be randomly chosen for payment and subjects did not receive information about the outcomes of previous rounds.

Although Bayesian learning effects cannot enter the series of single-shot games, we can probably detect a different form of learning. As experimental time passes, subjects might understand the structure of the game better and can also gather more routine in making decisions. Previous experiments of iterated PD, public good, and IPG games found that subjects approach the all-defection equilibrium over time (Isaac, McCue, and Plott, 1985; Andreoni, 1988; Andreoni and Miller, 1993; Bornstein, Winter, and Goren, 1996; Goren and Bornstein, 2000), which results in decreasing cooperation rates. Besides learning the structure of the game over time, another possible explanation is offered by Kreps et al. (1982). If subjects have an initial belief about others being irrational, it can be better for them to contribute initially, even if they are completely rational. Later on, they update their beliefs and they are then less likely to contribute. Experiments that included a “coincidental” restart of the game, however, found that contribution rates after restart went back up to virtually the same contribution rates as observed in the very first game (Andreoni, 1988). Based on these results Dawes and Thaler (1988) refuted the learning and the belief-updating hypotheses. However, by considering internalized reputation effects and image scoring (Wedekind and Milinski, 2000; Bienenstock, 2001) that can be mistakenly activated in single-shot games, we could explain decreasing contribution rates after every restart of repeated games in combination with the learning hypothesis.

If contribution propensities are not stable in the single-shot games within experimental parts, then we have to include an independent learning effect in the analysis and we have to drop our assumption that intra-individual variation (ξ_{ir}) has a zero expected value. As parts were separated by breaks, instead of checking for a single learning trend, it is better to distinguish between a within part and a between part learning trend in the analysis.

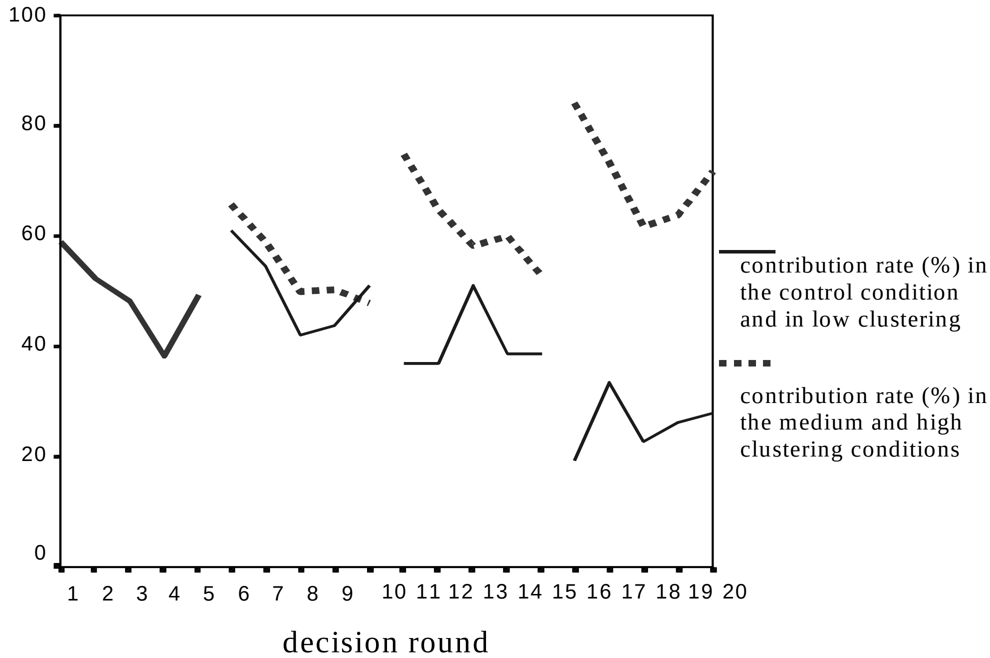

Figure 3.6.5.1 Contribution rates by decision rounds

Note: Part I: rounds 1-5; Part II: rounds 6-10; Part III: rounds 11-15; Part IV: rounds 16-20.

We found a decay of contribution for the series of single-shot games in our experiments (see Figure 3.6.5.1). Contribution rates decreased for those, who had some misunderstanding of the task before the game, but also for those, who answered quiz questions correctly. Besides the decreasing within part trend, in the last round of every part contribution rates increased significantly. It is a quite surprising result, since subjects knew that the outcome of the last round would neither be announced. This is exactly the opposite of what would be predicted on the basis of arguments of traditional game theory even if subjects had the incorrect perception that they are playing repeated games. This *endgame effect* was there independent of gender, social orientations, previous decisions, and other personal characteristics. By analyzing last rounds only, model parameters were similar to those values that were reported in Table 3.6.4.1, including the insignificant effect of internalized selective incentives. It means that higher contribution propensities in the last rounds cannot be explained by the reduction of cognitive dissonance ("in the last round I have to be nice, otherwise I cannot look at my fellow neighbors"). The resulting U-shape trend, however, has some correspondence to experimental findings in the iterated two-person PD and in collective action games (Rapoport and Chammah, 1965; Guttman, 1986). In our experiments, linear trend estimation provides the following parameters for within part learning:

$$P_{ri} = \ln\left(\frac{P_{ri}(C)}{P_{ri}(D)}\right) = 0.516(0.084) - 0.207(0.032)r + 0.376(0.112)r^5 + \varepsilon_i, \quad (3.6.5.1)$$

where r denotes the decision round number within the part, r^5 is a dummy for the last round in each part, and standard errors are in parentheses. Multilevel analysis gives an estimate of 0.693 (S.E.=0.089) for inter-individual variance. All effects are significant at the 0.1% level. The constant shows the average predicted contribution propensity for the first decision. We also tried to fit a non-linear model for within part learning, but goodness of fit did not increase significantly. Hence we included a linear trend element and an endgame effect as control variables in the model explaining contribution propensities (cf. Table 3.6.5.1). Model parameters are quite close to those in equation (3.6.5.1).

For between parts learning let us first report results from the low clustering condition (lower curve on Figure 3.6.5.1). In this condition no direct social control was introduced in between the parts, therefore Parts II, III, and IV were identical. Linear curve estimation provided a parameter of –0.5257 (S.E.=0.089, $p<0.001$) for between parts learning. It means that estimates of contribution rates in the low clustering conditions are 64.09%, 51.34%, 38.41%, and 26.94% in the four different parts, respectively.

We included both types of trend elements in the models in Table 3.6.5.1 as control variables. They were highly significant, as well as the puzzling endgame effect. But learning trends and endgame effects were not the only unexpected procedure effects. After controlling for the result of the last repeated game (if there was any), it emerged that a mutually harmful draw (clash punishment) “burns in” the memory of subjects and increases contribution propensities also in the single-shot games. This is probably an indication of that subjects use their long-term memory to estimate whether or not their decision could make a difference for the outcome in the forthcoming single-shot game. If they believe that a draw will occur, a single individual contribution can turn the outcome to victory. We did not anticipate these disturbances, but as we control for them in the analysis, results show the real effect of the main variables.

There are remarkable changes in the parameter estimates of social control. The effect of internalized selective incentives became significant and the significant effect of internalized traitor rewards has disappeared. The large increase in the estimate of baseline contribution propensity (constant) also indicates that the omission of learning trends resulted in a systematic bias in previous estimates in Table 3.6.3.1 and 3.6.4.1. Because of the negative between parts tendency, the baseline contribution rate was underestimated and the decrease between Part I and Part II was attributed to the effect of internalized traitor rewards. On the basis of the analysis reported in Table 3.6.5.1, we have to conclude that after controlling for a negative learning tendency, on average, traitor incentives in an internalized form do not influence the decision of subjects. On the other hand, we have to be careful with this interpretation and also with the confirmation of the existence of internalized selective incentives. The inclusion of a between parts trend in a linear functional form in the analysis does not stand on a firm theoretical basis. Furthermore, since the high correlation with experimental manipulations (the introduction of minimal contact and monetary forms of social control), the learning effect might include part of influence that we should attribute to other variables.

Table 3.6.5.1.a. Results of multilevel logistic regression on contribution propensities with personal characteristics and procedure effects

independent variable	hypothesis about the direction of effect	multilevel model with fixed slopes of main effects	multilevel model with random slopes of main effects
FIXED EFFECTS			
Main variables			
α (constant) baseline contr. propensity	?	1.378** (.423)	1.516*** (.409)
s_0 internalized selective incentives	+	.186* (.082)	.188* (.081)
s direct selective incentives	+	.769*** (.109)	.699*** (.127)
b_0 internalized behavioral confirmation	+	.586*** (.067)	.591*** (.080)
b direct behavioral confirmation	+	.718*** (.108)	.705*** (.126)
t_0 internalized traitor rewards	-	.165 (.086)	.142 (.086)
Personal characteristics and other subject-level variables			
gender (1=male)		-.176 (.143)	-.196 (.137)
student at the university (1=yes)		-.219 (.370)	-.352 (.357)
studies at the law faculty		-.109 (.366)	-.015 (.351)
studies natural sciences		-.057 (.344)	-.065 (.330)
studies economic, business, or spatial science		-.030 (.335)	.095 (.322)
studies social sciences		.068 (.309)	.136 (.296)
student of literary studies or arts		.056 (.316)	.133 (.303)
did a similar experiment before		-.154 (.136)	-.188 (.131)
strong risk aversion towards gains		-.163 (.135)	-.180 (.129)
strong loss aversion		.115 (.134)	.132 (.128)
consistent answers on social orientation questions		-.374* (.181)	-.400* (.173)
prosocial orientation		.511** (.183)	.487** (.175)
egalitarian orientation		.388* (.176)	.392* (.169)
number of acquainted subjects in the experiment		-.079 (.088)	-.093 (.085)
delay (minutes) at the start of the experiment		.008 (.007)	.006 (.007)
quiz questions answered correctly %		-.005 (.004)	-.005 (.004)
Procedure effects			
within part trend		-.215*** (.036)	-.213*** (.036)
endgame effect		.373** (.125)	.370** (.126)
between parts trend		-.397*** (.060)	-.379*** (.061)
last iterated game was a draw		.538*** (.149)	.515*** (.152)
last iterated game was lost		.185 (.122)	.199 (.125)
last iterated game was won		.214 (.123)	.275* (.125)

cont. next page

RANDOM EFFECTS		
interindividual variance σ^2	$.574^{+++}$ (.083)	$.559^{+++}$ (.116)
σ^2_{ui} (s_0)		.000 (.000)
σ^2_{ui} (s)		$.322^{+++}$ (.152)
σ^2_{ui} (b_0)		$.202^{+++}$ (.096)
σ^2_{ui} (b)		$.421^{+++}$ (.246)
σ^2_{ui} (t_0)		.002 (.050)
Covariances are reported below		+
-2 Log Likelihood model	4247	4198
Improvement χ^2 (df) for model in right column		49*** (20)
vs. previous model	183*** (6)	184*** (6)

Table 3.6.5.1.b. Random effects: estimated covariances

σ_{uxy}	ε_i	s_0	s	b_0	b
s_0	.000 (.000)				
s	-.163 (.109)	.000 (.000)			
b_0	.037 (.083)	.000 (.000)	$-.192^{+}$ (.090)		
b	$-.287^{+}$ (.133)	.000 (.000)	.063 (.143)	-.084 (.123)	
t_0	-.018 (.071)	.000 (.000)	.476 (.169)	-.054 (.117)	.152 (.180)

Notes: N=4011 decisions for 203 subjects. Iterative Generalized Least Squares estimates. Numbers in parentheses are standard errors. * significant at the 5% level, ** significant at the 1% level, ***significant at the 0.1% level (two-tailed).
For testing random effects it is more appropriate to use deviance tests: $^{+}$ significant at the 5% level, $^{+++}$significant at the 0.1% level (significance of difference in deviance compared to model without random slopes, for random covariates deviance is compared to model without random covariates).

There is another complication in relation to the difference in contribution propensities between Parts I and II. *Silent identification* (Bohnet and Frey, 1999) enters social dilemma experiments, when subjects are able to see each other. The visibility of others decreases social distance, allows for empathy and helps to conceptualize the experimental situation. We might predict this to be correlated with utility weights for social identity, consequently with higher contribution propensities. However, we cannot separate this effect from the influence of internalized social incentives that are not contingent on predictions (selective incentives and traitor rewards). If silent identification is a valid mechanism in the IPG game, our analysis overestimates the effect of internalized selective incentives. The unexpected positive sign of the t_0 estimate can also be partly explained by silent identification.

3.6.6 Interaction effects

As we have seen in Table 3.6.5.1, the significant effect of internalized traitor rewards disappeared after the inclusion of learning trends. It might be possible that we wrongly conceptualized this form of social control and traitor rewards could be there, but have another nature. They might stem from the presence of the other group as a whole or they exist only in certain dyadic relations.

The extension of the model by interaction effects helps with some clarification (see Table 3.6.6.1). It seems that internalized traitor rewards are activated in the dyadic context, but not in every neighborhood relation. Only neighbors of the opposite sex provide a significant control in the form of traitor rewards. This indicates that internalized pressure against contribution in the presence of opposite group members is activated only, if substantive distinction can be made apart from minimal group membership. Gender is possibly the most apparent characteristic that can be the source of this distinction between strangers. With respect to the interaction between gender and internalized behavioral confirmation, we did not find a significant effect on contribution propensities. However, descriptive statistics showed that subjects expected contribution more from fellow neighbors of the same sex and additionally, women were expected to contribute more.

Acquainted neighbors did not experience stronger social control than unknown ones did. Similar to the insignificant effect of the number of acquainted subjects in the experiment, this result can be probably attributed to the fact that they were not close acquaintances or to subjects' tendency to view laboratory conditions as impersonal. Prosocial and egalitarian attitudes were not correlated with higher relative weight of internalized social control. Only the interaction between traitor rewards and prosocial orientation proved to be significant. This effect indicates that subjects with prosocial orientation like to be "local heroes", who contribute even if they are surrounded by members of the other group. This is another indication of how *prosocial attitudes can be harmful in the intergroup context*.

Table 3.6.6.1.a. Results of multilevel logistic regression on contribution propensities with personal characteristics, procedure effects, and cross-level interactions

independent variable	hypothesis about the direction of effect	model with fixed slopes	model with random slopes
FIXED EFFECTS			
Main variables			
α (constant) baseline contr. propensity	?	1.346*** (.402)	1.491** (.477)
s_0 internalized selective incentives	+	.176* (.082)	.165* (.084)
s direct selective incentives	+	.769*** (.110)	.745*** (.135)
b_0 internalized behavioral confirmation	+	.589*** (.119)	.618*** (.141)
b direct behavioral confirmation	+	.703*** (.109)	.681*** (.125)
t_0 internalized traitor rewards	-	.223 (.132)	.238 (.134)
Personal characteristics and other subject-level variables			
gender (1=male)		-.089 (.146)	-.135 (.143)
student at the university (1=yes)		-.177 (.372)	-.201 (.364)
studies at the law faculty		-.162 (.368)	-.136 (.360)

cont. next page

studies natural sciences	-.101 (.349)	-.161 (.341)
studies economic, business, or spatial sciences	-.080 (.339)	-.002 (.330)
studies social sciences	-.001 (.312)	.000 (.305)
student of literary studies or arts	.045 (.317)	.066 (.309)
did a similar experiment before	-.179 (.136)	-.221 (.133)
strong risk aversion towards gains	-.172 (.134)	-.157 (.132)
strong loss aversion	.131 (.133)	.164 (.131)
consistent answers on social orientation questions	-.397* (.180)	-.404* (.176)
prosocial orientation	.330 (.206)	.353 (.202)
egalitarian orientation	.419* (.203)	.394* (.200)
number of acquainted subjects in the experiment	-.066 (.089)	-.066 (.087)
delay (minutes) at the start of the experiment	.006 (.007)	.006 (.007)
quiz questions answered correctly %	-.004 (.005)	-.005 (.005)
Procedure effects		
within part trend	-.178 (.121)	-.188 (.122)
endgame effect	.379** (.126)	.381** (.127)
between parts trend	-.397*** (.061)	-.386*** (.062)
last iterated game was a draw	.527*** (.150)	.495** (.157)
last iterated game was lost	.180 (.123)	.186 (.128)
last iterated game was won	.214 (.124)	.266* (.128)
Cross-level interactions		
$t_0\times$ number of acquainted opposite neighbors	-.153 (.196)	-.164 (.194)
$b_0\times$ number of acquainted fellow neighbors	.302 (.261)	.338 (.312)
$t_0\times$ number of opposite neighbors of the other sex	-.351** (.134)	-.373** (.137)
$t_0\times$ number of male opposite neighbors	.191 (.134)	.156 (.136)
$b_0\times$ number of fellow neighbors of the same sex	-.038 (.084)	-.128 (.102)
$b_0\times$ number of female fellow neighbors	.302 (.261)	.017 (.108)
$t_0\times$ prosocial orientation	.275* (.131)	.256* (.132)
$b_0\times$ prosocial orientation	.052 (.134)	.098 (.161)
$t_0\times$ egalitarian orientation	-.057 (.149)	-.025 (.149)
$b_0\times$ egalitarian orientation	.039 (.143)	.004 (.172)
within part trend $\times$ quiz questions correct %	.000 (.001)	.000 (.001)
RANDOM EFFECTS		
interindividual variance σ^2	$.563^{+++}$ (.082)	$.512^{+++}$ (.084)
σ^2_{ui} (s_0)		.000 (.000)
σ^2_{ui} (s)		$.549^{+++}$ (.187)
σ^2_{ui} (b_0)		$.143^{+++}$ (.089)
σ^2_{ui} (b)		$.379^{+++}$ (.240)
σ^2_{ui} (t_0)		.000 (.000)
Covariances are reported below		+
-2 Log Likelihood model	4211	4169
Improvement χ^2 (df) for model in right column		42** (20)
vs. previous model	36*** (11)	29** (11)

Table 3.6.6.1.b. Random effects: estimated covariances

σ_{uxy}	ε_i	s_0	s	b_0	b
s_0	.000 (.000)				
s	.037 (.107)	.000 (.000)			
b_0	.004 (.072)	.000 (.000)	-.145 (.093)		
b	-.200^{++} (.118)	.000 (.000)	.201 (.152)	-.031 (.116)	
t_0	.000 (.000)	.000 (.000)	.000 (.000)	.000 (.000)	.000 (.000)

Notes: N=4011 decisions for 203 subjects. Iterative Generalized Least Squares estimates. Numbers in parentheses are standard errors. * significant at the 5% level, ** significant at the 1% level, ***significant at the 0.1% level (two-tailed).
For testing random effects it is more appropriate to use deviance tests: $^{++}$ significant at the 1% level, $^{+++}$significant at the 0.1% level (significance of difference in deviance compared to model without random slopes, for random covariates deviance is compared to model without random covariates).

Additionally, we included an interaction variable to test whether or not subjects, who did not fully understand the experimental task, have different learning tendencies. The insignificant effect indicates that learning the structure of the game during the experiment is a general tendency and does not depend on the initial stage of understanding.

3.7 Discussion

The focus of this chapter differed from the mainstream experimental tradition of social dilemmas and attempted to incorporate sociological insights into the explanation of individual contribution rates. Our main objective was to show how internalized and direct social control enter into simple experimental situations and can affect individual decisions in an intergroup competition situation that we modeled by an Intergroup Public Goods game (Rapoport and Bornstein, 1987). As an aggregated result of different forms of social control, we tried to demonstrate *why segregation might induce the emergence of conflict between groups*. To discover the underlying mechanisms, we investigated what is the exact nature of *social control* and what are the forms that are already present in a condition with only eye contact between subjects. In order to test our hypotheses we introduced a *unique experimental design* based on special seating arrangements in the laboratory. With our setup we were able to target social control of the immediate environment that is believed to be influential also in real life.

By comparing clustering conditions we found that intergroup conflict was least likely in a completely mixed setting and was most likely when members of the groups were seated according to a segregated pattern, which confirms the segregation hypothesis. Furthermore, as predicted, the segregation effect was stronger under normative pressure than in the confirmation pressure condition.

By analyzing individual decisions, we uncovered mechanisms of social control that cause the segregation effect at the aggregated level. We found that *behavioral confirmation* is the form of social control, which strongly affects individual contribution propensities, also in an internalized form. Subjects adjusted their decisions towards the expected decision of their fellow neighbors even when eye contact is established among them. Our model parameters indicate that under the chosen reward structure, internalized confirmation pressure affected contribution propensities as much as monetary confirmation incentives did. We did not find, however, strong support for the presence of other forms of internalized social control in the minimal contact condition. Internalized selective incentives continued to have a significant effect after controlling for a between parts trend. However, the net effect of these incentives might diminish, if we consider silent identification (Bohnet and Frey, 1999) as a disturbing factor. Internalized traitor rewards might be activated in a dyad with minimal contact, but it is not a general mechanism. We found its clear presence only between neighbors of the opposite sex. Direct social control that was introduced in a form of additional monetary incentives had a significant effect.

In contradiction to what we predicted in the segregation hypothesis, contribution rates in the minimal contact condition were highest in the medium clustering condition. A possible explanation is that there is a *ceiling effect* similar to what we observed in high ranges of clustering in the simulation results (cf. Figures 2.8.1.2, 2.8.2.1, 2.8.2.2). This explanation is supported by evidence of high likelihood of conflict in the medium clustering condition (cf. Table 3.6.1.1). Another reason might be that the strength of internalized social control is a nonlinear function of the number of fellow neighbors. As a consequence, there is a marginal decrease in the segregation effect on the likelihood of intergroup conflict and already medium levels of clustering are often associated with harmful outcomes.

Among personal characteristics, we found a strong predicting power of social orientations. Subjects with prosocial and egalitarian attitudes found to be more contributive and consequently were found to be also more responsible for the emergence of mutually harmful outcomes between the groups. Another indication of that prosocial orientations are correlated with more generous behavior for the in-group, but more hostile behavior towards the out-group, is the positive interaction effect of traitor rewards and prosocial orientation. This implies that subjects with prosocial orientation behave as local heroes. If they are surrounded by members of the other group, they do not surrender at all. As a macro consequence, mutually harmful outcomes can occur even in the case of complete mixing, if there are enough prosocial individuals.

Although these effects appeared to be quite strong, we have to mention a limitation of the model. Little attention was devoted to individual utility functions and we were only interested in general tendencies. A complex utility assessment, however, is usually beyond consideration because it requires a lot of experimental time and answers are often inconsistent.

The analysis of the repeated IPG games included in these experiments is a target of subsequent chapters. First, in Chapter 4 we introduce theoretical concepts about effects of temporal embeddedness that we predict to influence repeated intergroup competition.

Appendix

3.A.1 Instructions in the experiment

Instructions in the experiment were translated into Dutch. The first pages ("Instructions" and "What do you do in this experiment") were also distributed on paper.

Instructions at the arrival to the laboratory

Team competition: An experiment
Instructions

- After you have listened to the verbal introduction and read the instructions, there will be time to ask questions.
- The instructions on your screen and on the paper are exactly the same.
- Please leave the instructions next to your computer when leaving the room.
- VERY IMPORTANT: Any kind of communication with other participants is *strictly prohibited* during the experiment. We kindly ask you to keep this rule. Violation can result in money withdrawal from your earnings or your exclusion from the experiment.
- The experiment will take approximately 90 minutes and you will receive your payment at the end of the experiment. How much you earn will depend on your decisions and on the decision of other participants. This will be explained in detail on the next page.

Instructions at the start of the experiment

What do you do in this experiment?

This experiment is about competition between groups. There are 10 participants in this experiment; you are one of them. The participants are divided into ***two teams***. Both teams consist of 5 members. Your group membership and the membership of other participants will be announced later in the experiment.

The experiment will take several decision rounds. All decision rounds will count in your final payment! At the end of the experiment, the computer will calculate your AVERAGE earning in the decision rounds. This will be your real money, what you will receive in cash after the experiment has finished. If you want to earn a lot of money, try to gather as much symbolic money as possible!

At the beginning of each decision round, you will receive 11 Dutch guilders as a bonus. The bonus appears on your screen and you have to decide

- whether to give away this bonus to help your team to reach success in the team competition
- or to keep the bonus.

Other participants have to make exactly the same decisions. Decisions are kept strictly confidential. You will not know the decision of others and they will not know your decision.

The team, in which more members gave away their bonuses, wins the competition and the other team loses. For every member in the winning team (irrespective of his or her own decision) the same amount of

15 guilders are distributed. From every member of the losing team (irrespective of his or her own decision) 15 guilders are subtracted. In the case of a draw, 11 guilders are subtracted from every participant (irrespective of his or her own decision). The only exception is when in both teams there are *less, than 3 individuals* who are willing to give away their bonuses. In this case not enough interest is recorded in the team competition and therefore no money is distributed or subtracted.

For your better understanding of how much money you earn in one round please consider the following table:

Table 1 Your earnings (Dutch guilders) in one round of the experiment

	result of the competition	*not enough interest*	*your team wins*	*draw*	*your team loses*
your decision	**give away your bonus to help your team**	0	15	-11	-15
	keep your bonus	11	11+15 = 26	-11+11 = 0	-15+11 = -4

After the experiment, you will receive the average of what you have earned in the different rounds. This amount will be increased by 15 guilders to guarantee you a positive earning regardless of what you do.

Please remember that all team members are real people. The computers are only here to record your decisions. Your decisions will be correctly saved and you will always receive completely true information during the experiment.

Before the real experiment starts, we would like to be certain that you have understood the decision task correctly. You will be kindly asked to answer the quiz on your computer. In this quiz of understanding, the decision of others is generated randomly by the computer. The fictive money you gain in the quiz will not count in your real payment.

Now you can ask your questions.
Good luck!

Instructions at the start of the quiz

As a warming-up, you will receive some questions that help you to understand the decision task correctly.

We would like to be certain that you understood the decision task correctly. For this reason please consider the following possible outcomes. We ask you to select what you think your reward would be in the given situation. Now and also later during the experiment you can get help by looking at Table 1 in the instructions.

Instructions at the start of Part I

The quiz hopefully helped you to understand the decision task correctly. Once again, during the experiment you can always get help by looking at Table 1 in the instructions.

This is the real experiment. What you gain in this round, will count in your score. Be careful, if you run out of time, the computer will decide randomly for you. For each case that this happens there will be 1% decrease in your final payment.

We will also ask you to give some predictions about the outcome of the competition. Please take into consideration, that every prediction that you do not answer, decreases your final payment by 0.2%.

From the first 5 rounds, only one will count in your final payment. The computer will select this round randomly.

We kindly ask you to pay attention to the order of answers, because this changes during the experiment.

We kindly ask you to avoid any form of communication with other participants.

As soon as this information disappears from your screen, the experiment starts.

Instructions at the start of Part II

Announcement of group membership

Now we announce your team membership and the membership of others. The members of the two teams will receive red and green flags. Participants with red flags consist of one team and participants with green flags consist of the other team.

Now the experimenter distributes the flags. Please look around you and see whether your neighbors are from your team or from the other team. We kindly ask you not to stand up from your place and not to talk with other participants. Please do not remove your flag from your computer during the experiment.

You can also see who are your neighbors from your screen. Be aware that your neighbors know in which group you are and they might know what your decision was.

Now we will continue with the experiment. You will receive 11 Dutch guilders as a bonus at the beginning of each decision round. As before, you will be asked to decide

- whether to give away this bonus to help your team to reach success in the team competition
- or to keep the bonus.

Good luck!

Instructions at the start of Part III

Modifications

You are still in the same group as before and your teammates did not change either. At the beginning of the following decision rounds, you will receive 11 Dutch guilders as a bonus, as before. You again have to decide

- whether to give away this bonus to help your team to reach success in the team competition
- or to keep the bonus.

Other participants have to make exactly the same decision. Just to recall how much money you can earn in one round please consider the following table again:

Table 1 Your earnings (Dutch guilders) in one round of the experiment

	result of the competition	*not enough interest*	*your team wins*	*draw*	*your team loses*
your decision	**give away your bonus to help your team**	0	15	-11	-15
	keep your bonus	11	11+15 = 26	-11+11 = 0	-15+11 = -4

This is the same procedure as before, but now we introduce some modifications about the payments.

Please look around whether your neighbor is from your team. A neighbor is a participant, who is sitting next to you on your left and/or on your right. Your neighbor is from your team if there is a flag of the same color on your monitor and on his or her monitor. *If you do not have a neighbor from your team, then there is no modification in the method of payment.*

[for conditions in which behavioral confirmation is introduced]

You will earn 5 guilders extra in every round, if you do the same as your neighbor in the same round.

If you have two neighbors from your team, this rule applies to both of them. It means that you can earn 10 (2x5) guilders if you and both of your neighbors do the same in that round. If the decision of these two neighbors is different, then you earn 5 guilders.

[for conditions in which selective incentives are introduced]

You will earn 5 guilders extra, if you have one neighbor from your team and you decide to give away your bonus. You will earn 10 (2x5) guilders, if you have two neighbors from your team and you decide to give away your bonus.

[for conditions in which selective incentives or behavioral confirmation is introduced]

Everyone, who has neighbors from his or her team, has this opportunity to earn extra money.

Now you can ask questions.
Good luck!

Instructions before Part IV were constructed similarly.

3.A.2 Measurement of risk attitudes

Experimental evidence shows that individuals have different risk attitudes towards gains, losses, and mixed rewards, and typical risk preferences are best described by an S-shape function (Tversky and Kahnemann, 1992). Since we had to avoid a long questionnaire, we could only make a very superficial assessment of individual attitudes towards risk. We constructed six questions, enough to determine some basic characteristics of utility functions, but not enough to reconstruct them entirely.

As a measurement, we used the method of preference comparisons (see Farquhar, 1984). In three questions standard gambles and in three questions complex gambles were applied. In both cases, separate questions were related to gains, to losses, and to mixtures of gains and losses. With the standard gamble method we measured whether the subject is characterized as risk-averse or risk-seeking. The complex gamble was used to get an indication of individual behavior toward compound lotteries. Since decision in the IPG game has some correspondence to such a compound lottery, this instrument probably measures risk attitudes in this specific context more accurately. The gambles were constructed to represent the possible decision dilemmas subjects have faced during the experiment. Therefore, numbers in the gamble were close to monetary rewards that were at stake in the IPG game (see Table 3.A.2.1).

Table 3.A.2.1 Measurement of risk preferences

method	*range*	*option 1*	*option 2*
standard gamble method	gains	*f* 26	[*p* 0.5, *f* 15; *p* 0.5, *f* 37]
	losses	- *f* 26	[*p* 0.5, - *f* 15; *p* 0.5, - *f* 37]
	gains and losses	*f* 0	[*p* 0.5, - *f* 11; *p* 0.5, *f* 11]
complex gamble	gains	[*p* 0.5, *f* 26; *p* 0.5, *f* 31]	[*p* 0.5, *f* 15; *p* 0.5, *f* 37] + [*p* 0.5, *f* 0; *p* 0.5, *f* 5]
	losses	[*p* 0.5, - *f* 26; *p* 0.5, - *f* 31]	[*p* 0.5, - *f* 15; *p* 0.5, - *f* 37] + [*p* 0.5, *f* 0; *p* 0.5, - *f* 5]
	gains and losses	[*p* 0.5, *f* 0; *p* 0.5, *f* 5]	[*p* 0.5, - *f* 11; *p* 0.5, *f* 11] + [*p* 0.5, *f* 0; *p* 0.5, *f* 5]

We classified subjects as strong risk-averse towards gains, if they chose the risk-averse option (option 1 in Table 3.A.2.1) both in the standard and in the complex gamble with positive rewards. They were classified as strong loss-averse, if they chose the risk-seeking option (option 2 in Table 3.A.2.1) both in the standard and in the complex gamble with negative rewards. These variables were used as predictors in the analysis of contribution propensities (Sections 3.6.4 till 3.6.6).

3.A.3 Measurement of social orientations

Although in theory many types of social values are identified (cf. Liebrand, 1986), experimental evidence shows that the vast majority of people can be characterized as prosocial (cooperative), individualistic, or competitive (Suleiman and Or-Chen, 1999). For the classification of individuals into these categories two techniques are used. The *triple-dominance method* consists of a series of decomposed games with an unknown, randomly assigned person (see van Lange et al., 1997). Individuals have to make decisions between alternatives of paired monetary rewards for themselves and for the other. The triple-dominance measure of social values has the clear advantage that from few (6 or 9) questions social orientations can be detected in a reliable way. In comparison, the *ring measure* of social values (McClintock and Liebrand, 1988; van

Lange, 1999) consists of choices between 24 pairs of self-other outcome combinations. The ring measure gives a better approximation of the weights assigned to the outcome for the self and for the other. In the self-other outcome space, these weights can be conceptualized in terms of the angle θ between the horizontal axis and the line connecting the individual's ideal point with the origo (Doi, 1994):

$$SU_i = x\cos\theta_i + y\sin\theta_i, \tag{3.A.3.1}$$

where x denotes the outcome for the self and y denotes the outcome for the other. A graphical representation in Figure 3.A.3.1 indicates ranges of prosocial (cooperative), individualistic, and competitive orientations. Dashed lines separate these ranges.

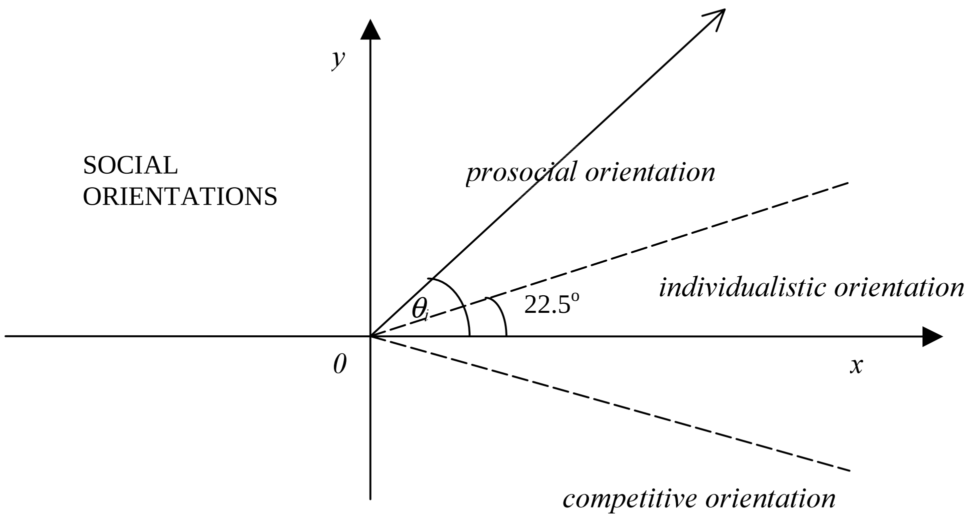

Figure 3.A.3.1. Graphical representation of social orientations with an example of prosocial (cooperative) individual values

Due to this model of social values, individuals have linear indifference curves in the self-other outcome space, with a slope of

$$RS_i = \frac{\cos\theta_i}{-\sin\theta_i}, \tag{3.A.3.2}$$

indicating the constant rate of substitution for individual *i.* For individuals, who are exactly at the border of individualism and competition, the rate of substitution is 1+√2. It means that they enjoy the same social utility, when one unit decrease in their own outcome is compensated by 1+√2 unit decrease in the outcome of the other. For competitive individuals, the substitution rate is higher than this value. Individuals at the border of cooperation and individualism have a -1-√2 substitution rate. They have the same social utility; when one unit decrease in their own outcome is compensated by

$1+\sqrt{2}$ unit increase in the outcome of the other. For individuals with cooperative social orientation, less compensation is satisfactory.

The triple-dominance measure of social orientations revealed good internal consistency and test-retest reliability over a considerably long period (e.g., Kuhlman, Camac, and Cunha, 1986; van Lange et al., 1997). However, the method is constructed only to a specific range of possible outcomes. All outcomes used by the measurement are expressed as gains, both for the self and for the other (see van Lange et al., 1997). Furthermore, the outcome for the self is never smaller than the outcome for the other.

Because of these restrictions, some paradoxes might arise showing inconsistencies and incompleteness of the social utility function in equation (3.A.3.1). Since individual behavior is significantly different toward gains and losses for the self (Tversky and Kahneman, 1986), we expect violations of the consistency of social orientations, where subjects have to decide over self-other pairs of losses or mixtures of gains and losses. Violations could also occur due to the decreasing marginal utility of money. Such a phenomenon might also exist for the outcome of the other. Furthermore, individual decisions can differ when a higher outcome is offered for the other from when it is offered to self. It is well known in social psychology that relational interpersonal comparisons affect psychological well-being (utility).

In a recent study, van Lange (1999) proposed an integrative model of social orientations. He claimed that the social utility function of cooperative individuals includes a weight assigned for the equality of the outcomes. His experimental findings provided support for this hypothesis. Previously, Knight and Dubro (1984) incorporated egalitarianism into the social utility function. In the analysis of individual behavior in the ultimatum game, which has a close correspondence with the measurement of social orientations, fairness was a major motivational concern for non-selfish behavior (Bethwaite and Tompkinson, 1996). Fairness was conceptualized as a utility element for the distance between individual payoffs. The smaller the difference is the higher the social utility of the individual.

The difference between the outcomes for the self and for the other may also matter for individuals with competitive or individualistic orientations. Kuhlman, Camac, and Cunha (1986) discussed that competitive individuals are either appetitive or aversive. Appetitive individuals have a positive desire to surpass others. On the other hand, aversive competitiveness results from a fear of being surpassed. Bethwaite and Tompkinson (1996) introduced this concept into the utility function as envy. Their experimental results showed that it was less important than fairness, but still it had an influence on behavior in ultimatum games (Bethwaite and Tompkinson, 1996).

For the problems with social orientations that occur as anomalies Schulz and May (1989) provided an elegant solution. There are individuals, who are individualistic when the outcome for self is smaller than it is for other and they are cooperative when the outcome for self is larger than the outcome for other. For them, the parameter of social orientations, which is represented as the angle θ_i (see Figure 3.A.3.1) is different for the ranges $x<y$ and $x>y$ (let us denote these by θ_i^A and by θ_i^B). Assuming that the utility function is continuous at the point, where $x=y$, this special case of conditional utility fits in the more general description of social utility (Schulz and May, 1989), given as

$$SU_i = a_i x + b_i y - c_i |x - y| \tag{3.A.3.3}$$

where a_i, b_i are the individual weights assigned to the outcome for self and for other, and c_i is an individual parameter assigned to egalitarian and other relational interpersonal tendencies. In our special case,

$$a_i = \frac{\cos\theta_i^A + \cos\theta_i^B}{2};$$

$$b_i = \frac{\sin\theta_i^A + \sin\theta_i^B}{2};$$

and $$c_i = \frac{\cos\theta_i^A - \cos\theta_i^B}{2} = \frac{\sin\theta_i^B - \sin\theta_i^A}{2}.$$

For the sake of interpersonal comparisons and graphical representation, Schulz and May (1989) introduced the restriction $a_i^2+b_i^2+c_i^2=1$ for the weights. In this way, social orientations can be represented as points on the surface of a sphere. Simple social orientations that can be described properly by the utility function in equation (3.A.3.1) are located on the equator plane. There is a third dimension that represents the importance of interpersonal comparisons. Poles of the sphere mark extremely high importance (positive or negative) for interpersonal comparisons. Schulz and May (1989) detected that social orientations of many people fall outside of the equatorial, and thus cannot be described simply as individualistic, cooperative, or competitive.

Our experiments are conducted on computers, which made it possible to ask subjects to choose self-other outcome combinations, conditional on their previous response. A new measurement is proposed here, which uses only four questions, can classify subjects into subcategories of social orientations, can detect egalitarianism, and includes a consistency check. This method, similar to the triple-dominance measure of van Lange et al. (1997), is only related to behavior towards positive outcomes in the $x>y$ range. However, the method could be used in the positive $y>x$ range (4 additional questions), towards negative outcomes (4+4 questions), and also towards mixed incentive structures (gains and losses).

Table 3.A.3.1 Measurement of social orientations for $x>y$

question 1				question 2				question 3				question 4			
ty	*slf*	*oth*	*ty*	*ty*	*slf*	*oth*	*ty*	*ty*	*slf*	*oth*	*ty*	*ty*	*slf*	*oth*	*ty*
	24	2	*c*		25	10	c^{+}		25	25	*e*		29	7	*c*
all	29	14	*i*	*c*	29	14	c^{0}	c^{+}	30	27		*all*	34	19	*i*
	26	21	*p*		33	20	c^{-}		29	28			31	27	*p*
					27	4	i^{+}		25	25	*e*				
				i	29	14	i^{0}	c^{0}	23	22					
					28	19	i^{-}		21	20					
					27	12	p^{-}		25	25	*e*				
				p	25	15	p^{0}	c^{-}	22	18					
					23	17	p^{+}		20	15					
									25	25	*e*				
								i^{+}	24	22					
									21	20					
									25	25	*e*				
								i^{0}	26	20					
									25	15					
									25	25	*e*				
								i^{-}	27	20					
									28	15					
									25	25	*e*				
								p^{-}	27	22					
									30	15					
									25	25	*e*				
								p^{0}	27	23					
									29	20					
									25	25	*e*				
								p^{+}	27	24					
									29	22					

Notes: c=competitive, *i*=individualistic, *p*=prosocial (cooperative), *e*=egalitarian orientations, *ty*=type, *slf*=reward for the self, *oth*=reward for the other.

In the first step, subjects are classified into the three main categories (cooperation, individualism, and competition) in the classical way. They have to choose between three alternatives of outcome pairs. Options are constructed in order to represent the borders between the categories (dashed lines on Figure 3.A.3.1). In the second step, the θ_i parameter is approximated, assuming that c_i is close to zero. The third step is constructed to detect egalitarianism. If subjects choose for the egalitarian choice in this question, the assumption of $c_i \approx 0$ will be violated (assuming that θ_i was approximated correctly). The fourth question is a consistency check. Numbers that are close to the rewards of the experiment were applied (cf. Tables 3.A.3.1 and 3.A.3.2).

Table 3.A.3.2 Measurement of social orientations for $x<y$

question 5				question 6				question 7				question 8			
ty	*slf*	*oth*	*ty*	*ty*	*slf*	*oth*	*ty*	*ty*	*slf*	*oth*	*ty*	*ty*	*slf*	*oth*	*ty*
	12	14	*c*		13	22	c^+		25	25	*e*		17	19	*c*
all	17	26	*i*	*c*	17	26	c^0	c^+	26	28		*all*	22	31	*i*
	14	33	*p*		21	32	c^-		28	30			19	38	*p*
					15	16	i^+		25	25	*e*				
				i	17	26	i^0	c^0	30	30.5					
					16	31	i^-		32	33					
					15	24	p^-		25	25	*e*				
				p	13	27	p^0	c^-	27	28					
					11	29	p^+		28	30					
									25	25	*e*				
								i^+	27	30					
									28	35					
									25	25	*e*				
								i^0	26	30					
									25	35					
									25	25	*e*				
								i^-	24	30					
									22	35					
									25	25	*e*				
								p^-	23	30					
									21	33					
									25	25	*e*				
								p^0	23	28					
									21	30					
									25	25	*e*				
								p^+	23	27					
									20	29					

Notes: *c*=competitive, *i*=individualistic, *p*=prosocial (cooperative), *e*=egalitarian orientations, *ty*=type, *slf*=reward for the self, *oth*=reward for the other.