Social identity and trust
– An experimental investigation –

Werner Güth[a], M. Vittoria Levati[a,*], Matteo Ploner[b]

[a] *Max Planck Institute of Economics, Strategic Interaction Group, Jena, Germany*

[b] *LEM, Sant'Anna School of Advanced Studies, Pisa, Italy*

Abstract

We experimentally examine how group identity affects trust behavior in an investment game. In one treatment, group identity is induced purely by minimal groups. In other treatments, group members are additionally related by outcome interdependence established in a prior public goods game. Moving from the standard investment game (where no group identity is prompted) to minimal group identity to two-dimensional group identity, we find no significant differences in trust decisions. However, trust is significantly and positively correlated with contribution decisions, suggesting that "social" trust is behaviorally important.

JEL Classification: C72; C92

Keywords: Experiment; Investment game; Trust; Group identity

*Corresponding author. Max Planck Institute of Economics, Kahlaische Str. 10, 07745 Jena, Germany. Tel.: +49 3641 686629; fax: +49 3641 686667.

E-mail address: levati@econ.mpg.de (M. V. Levati).

1 Introduction

Economic transactions often require trust since contracts are incomplete and proper dealings cannot be enforced legally. Arrow (1974) defines trust as “a lubricant of the social system”, and much research has accumulated across various academic fields linking trust with institutional efficiency and economic growth (cf., e.g., Putnam, 1993; Fukuyama, 1995; La Porta et al., 1997). Recently, Rothstein and Uslaner (2005) have made a clear distinction between social (or generalized) trust, which links us to people who are different from ourselves, and “particularized” trust, which is restricted to our own kind. While the former is found to be correlated with a number of normatively highly desirable variables (such as more voluntarily giving, more growth, better health, and lower crime figures), particularized trust may yield serious social problems leading ultimately to a so-called “social trap” (cf., Rothstein, 2005). On these grounds, it is worthwhile studying to what extent individual identification with a group can foster the propensity to trust others in the same group or in another, e.g., neighboring, group.

The idea that shared social identity may provide a basis for “depersonalized” (or particularized) trust has been advocated by Brewer (1981). Some empirical research supports this argument. Laboratory studies using the “minimal group paradigm” (cf., Tajfel, 1970), for instance, indicate that inducing experimentally a seemingly innocent ingroup-outgroup classification suffices to activate ingroup-favoring behavior in reward allocation: no matter how arbitrary the group distinction is, individuals are more likely to cooperate with ingroup than outgroup members (see Schopler and Insko, 1992, for a review).

According to some authors (e.g., Jin and Yamagishi, 1997), expectations of reciprocity are critical for the emergence of ingroup-favoring behavior: we naturally anticipate more contact with ingroup than with outgroup partners and tend to expect more reciprocity by the former. Thus, trust behavior in group settings must be understood as an economic decision about risk and benefit. On the other hand, social identity and categorization theorists claim that ingroup-favoring behavior reflects a motivation to create and maintain positive distinctiveness for the group with which the self is identified (cf., e.g., Kramer, 1991; Brown, 2000). From this perspective, in-group trust must be understood as a social decision. Both explanations of identity-based trust, although divergent, can account for personalized trust, i.e., more cooperation with ingroup members.

Plenty of experimental research on trust and trustworthiness in the economic domain has shown, however, that individuals are willing to trust and to honor trust (i.e., to reciprocate) even when no group identity is prompted.[1] This suggests that cooperative attitudes are proper to an individual, and people tend, in general, to have faith in others.[2] The present experimental study is designed to assess whether these cooperative attitudes can actually be manipulated by artificially establishing the status of being the member of one group rather than another.

For measuring intra as well as intergroup cooperation we rely on the investment game introduced by Berg et al. (1995). In a dyadic relationship, a trustor can send all, some, or none of her monetary endowment to a trustee, who receives three times the trustor's sacrifice, i.e., the social benefits of giving are greater than the social costs. After learning how much she has

[1]Experiments in which trust and reciprocity are important are provided, among others, by Berg et al. (1995); Ortmann et al. (2000); Cox (2004).

[2]Broadly speaking, there exist two schools of thought about the determinants of social trust and cooperative dispositions (cf., Delhey and Newton, 2003). The first claims that cooperative attitudes are an individual's own characteristics (e.g., Uslaner, 1999; 2000). The second argues that preferences towards cooperation are a consequence of trusting cultures or social and political institutions (e.g., Putnam, 2000).

got (and therefore about the trustor's sacrifice), the trustee chooses how much of the received money she wants to return to the trustor. In this setting, trust is highly profitable due to a multiplier of 3, but also very dangerous due to the one-shot nature of the game.

One of our experimental treatments resembles the minimal group paradigm by merely assigning symbols to participants, where the symbols for the sender (or trustor) and the recipient (or trustee) can be either equal or different. Because of the low emotional and strategic content of such symbols, this kind of group treatment appears like the weakest form of introducing "social distance" (see, e.g., Buchan and Croson, 2004; Buchan et al., 2006). The device we employ is uncommonly minimal.[3] Yet, it fully respects all features of standard "minimal groups", i.e., no similarities among group members, no face-to-face interaction, and no shared fate or prior experience with the category (cf., Brewer, 1979).

In general, most previous laboratory studies of the minimal group paradigm draw on settings different from the investment game. An exception is the cross-cultural experiment of Buchan et al. (2006), investigating how "social distance" – manipulated via minimal groups – affects trust. Their categorization consists in assigning participants to one of four color-coded groups, and letting each group engage in some type of non-relevant communication. Half the participants are then paired with a member of their discussion group, and the other half with a member of a different discussion group. While Americans display a strong ingroup-outgroup bias, participants from China do not. The authors attribute these results to the different cultural orientation across countries.

[3]Most previous studies divide subjects into groups based on their preferences for various forms of art or political parties (cf., Tajfel et al., 1971; Gaertner et al., 1999), on how they perform a given task (cf., Tajfel et al., 1971), or even on whether they communicate prior to interaction or not (cf., Orbell et al., 1988; Frey and Bohnet, 1997).

Aside from being restricted to one country only, our experiment differs from Buchan et al.'s study in inducing group identity without communication. Moreover, on account of Worchel et al. (2000)'s assertion that individual behavior is guided by *several* identity concerns acting simultaneously, some of our treatments trigger group identity in two ways: "mere labeling" and "shared interests" or, using the language of psychology, "common fate" (see, e.g., Hornsey et al., 2003, and the references therein). These treatments start out with a public goods game between two subgroups (or neighboring regions), each providing a local public good whose benefits favor one's own region but transcend local boundaries. Hence, "labeling" refers to the two subgroups or neighborhoods, and "shared interests" (or "common fate") to the strength of external effects measuring how much one is affected by what *another* individual chooses to do. In such an environment, by varying the positive spillovers between regions, we can manipulate the interests in the own subgroup as compared to the whole group. When spillovers are very high, the welfare of each subgroup strongly depends on the combined performance of both subgroups. Conversely, when spillovers almost vanish, the welfare of each subgroup depends mainly on its own performance. It is hypothesized that experiencing common fate adds genuine content to group membership beyond mere labeling (in political science, Singleton and Taylor, 1992, refer to such strengthening of social identity as "mutual vulnerability").[4]

This way of operationalizing common fate (i.e., in terms of shared out-

[4]It may be argued that any (though very small) spillover between subgroups may inspire a collective social identity transcending one's subgroup (see, e.g., Dayton-Johnson, 2003). To check whether this is indeed the case, one could have run a control experiment with two totally isolated subgroups. Yet, we preferred not to do so because, from a methodological perspective, it is far more preferable to compare treatments differing just in one numerical parameter than treatments requiring different verbal instructions. Moreover, in light of our results (showing that subgroups are actually relevant to individuals), such control experiment with no spillovers seems less important.

comes that flow on from performance) is rather common in the psychological literature, even though it follows different procedures (see, e.g., Rosenbaum et al., 1980; Gaertner et al., 1999; Hornsey et al., 2003).[5] In particular, the experimental paradigm used by psychologists begins by creating two distinct social categories to which participants are assigned. The two groups/categories are then separated in order to engage in discussion or work together as segregated social groups. At this point, the researchers create groups composed of representatives from each category. The newly formed groups are given a task requiring cooperative interaction to reach a common goal. Manipulation of common fate concerns how the newly formed groups distribute the rewards for the performed task.

Notwithstanding the different modus operandi, the present study induces social categories boundaries in a similar manner, i.e., via interdependent rewards for contribution to a public good. Will such a manipulation of common fate be effective in producing higher levels of social identity among the intended pairs? If so, will group identities shape the degree to which people trust? Or will individuals be willing to cooperate regardless of the extent of group boundaries? These are the main questions underlying our research. Investigating whether artificially inducing social boundaries allows for more ingroup trust (as suggested by the minimal group paradigm) should shed light on the behavioral relevance of particularized trust as opposed to social trust.

[5] There exist, however, experiments in which common fate is not framed in terms of performance (see, e.g., Kramer and Brewer, 1984).

2 The games

To address our major issues we rely on two different games. Berg et al.'s (1995) investment game is used to assess people's propensity to cooperate. A linear public goods game systematically varying outcome interdependence is used to manipulate common fate within and between groups.

2.1 The investment game

Let A and B be two interacting players, each endowed with 10 ECU (Experimental Currency Unit). Trustor A can send any integer amount x_A to B, with $0 \leq x_A \leq 10$. The amount sent is then tripled and received by trustee B, who can return any integer amount x_B to A, with $0 \leq x_B \leq 3x_A$. The payoffs are $\pi_A = 10 - x_A + x_B$ for A, and $\pi_B = 10 + 3x_A - x_B$ for B. If trustee B maximizes her own monetary payoff, she will keep everything, i.e., $x_B^* = 0$ for all $x_A \in \{0, 1, 2, \cdots, 10\}$. Anticipating this, trustor A will send nothing, i.e., $x_A^* = 0$. This yields the solution payoffs $\pi_A^* = \pi_B^* = 10$. On the other hand, positive investments by player A increase the sum of both payoffs, which is $20 + 2x_A$ and maximal for $x_A^+ = 10$. Thus, social efficiency requires A to send everything. In this case, $\pi_A^+ + \pi_B^+ = 40$ that can be freely allocated by choosing x_B appropriately.

2.2 The linear public goods game

We induce shared interests between trustor A and trustee B by means of *local* public goods (Güth et al., 2005). Let $N = \{1, \ldots, 4\}$ be a society with 4 individuals. The society is composed of two subgroups, X and Y, with members $i \in X = \{1, 2\}$ and $j \in Y = \{3, 4\}$. All four individuals receive an endowment of 10 ECU. Members i of subgroup X can contribute to public good C by choosing $c_i \in \{0, 1, \ldots, 10\}$. Members j of subgroup Y

can contribute to public good D by choosing $d_j \in \{0, 1, \ldots, 10\}$. The size of the two public goods, therefore, is $C = \sum_{i \in X} c_i$ and $D = \sum_{j \in Y} d_j$.

Each individual's payoff is linear in her own contribution and in both public goods, and is given by:

$$\pi_i = 10 - c_i + \alpha C + \beta D \qquad \forall i \in X, \tag{1}$$

$$\pi_j = 10 - d_j + \beta C + \alpha D \qquad \forall j \in Y \tag{2}$$

with $1 > \alpha > \beta > 0$. Due to these inequalities, X-members profit more from C and Y-members more from D although, due to $\beta > 0$, all individuals gain from both public goods. Because of $\alpha < 1$, each player would maximize her own payoff by not contributing. However, imposing the usual assumption of public goods games, namely $2(\alpha + \beta) > 1$, implies that all individuals gain by fully contributing. To render full contribution worthwhile even when being concerned with subgroup-efficiency only, one can require $2\alpha > 1$. In this case, there are already efficiency gains from a subgroup-perspective, regardless of what is expected from the other subgroup. By varying $\alpha - \beta$ (i.e., the tension between subgroups), we can manipulate the interest in the own subgroup/region as compared to the whole group/society.

3 Experimental treatments and procedures

Overall, we implement nine different treatments in a between-subjects design. The *control* treatment (hereafter C-treatment) is the standard investment game as described in Section 2.1.

The *labeling* treatments induce group identity in a minimal way. They differ from the C-treatment in that participants are assigned one of two different labels (either X or Y). As participants enter, they receive an envelop containing their label. Pairs with either the same or different labels then play the investment game. Participants are aware of their partner's label, i.e., they know whether the person they are matched with is carrying their same or a different label. Henceforth, *LAB-IN* will denote the treatment where the two interacting players have the same label, and *LAB-OUT* the treatment where they have different labels.

In the other six treatments the public goods game is played before the investment game so as to manipulate group identity through common fate. Like in the labeling treatments, half of the participants gets an X-label (the so-called X-types) and half a Y-label (the so-called Y-types). Groups with two X-types and two Y-types each then play the linear public goods game described in Section 2.2. Aside from assigning different labels, the public goods game relates players via positive spillovers affecting the X- and the Y-types differently. Subjects are informed – by payoffs (1) and (2) and a series of examples – that the externality generated by the player of the same type on one's payoff is greater than that generated by either player of the other type.

Intergroup tension is varied by distinguishing three treatments systematically varying the difference between α and β. In one treatment, with $\alpha = 0.55$ and $\beta = 0.45$, the difference between α and β is negligible and equal to $\epsilon = 0.1$.[6] In the intermediate treatment, with $\alpha = 0.65$ and $\beta = 0.35$, one has $\alpha - \beta = 3\epsilon = 0.3$. In the third treatment, with $\alpha = 0.95$ and $\beta = 0.05$, the difference $\alpha - \beta = 9\epsilon = 0.9$ is substantial. Henceforth, the three treatments will be addressed as ϵ, 3ϵ, and 9ϵ, respectively.

Whereas in case of $\alpha - \beta = \epsilon$ the tension between subgroups in minimal, when $\alpha - \beta = 9\epsilon$ the intergroup connection is rather weak compared to the

[6] We regard this treatment as an approximation of the minimal group design: though subjects are divided in two subgroups, spillovers between groups are very small. As already noted, this treatment is methodologically more proper than a treatment with no spillovers because the latter would have required different verbal instructions.

intense intragroup interest. The intermediate constellation with $\alpha - \beta = 3\epsilon$ allows us to check the monotonicity of possible $\alpha - \beta$-effects. The various marginal per capita returns are chosen so as to equalize the efficiency gains in all three treatments; specifically, $2(\alpha + \beta) = 2$ always. Furthermore, in all treatments, $2\alpha > 1$ meaning that perfect ingroup solidarity alone would inspire full contributions.

After the public goods game, participants play the investment game either with the member of the same subgroup/region (hereafter *IN*-treatment) or with a member of the other subgroup/neighboring region (hereafter *OUT*-treatment). The two matching procedures in the investment game and the three levels of outcome interdependence in the public goods game yield a 2×3-factorial between-subjects design. Table 1 summarizes the main characteristics of our nine treatments.

Insert Table 1 about here

The computerized experiment was conducted at the laboratory of the Max Planck Institute in Jena (Germany). The experiment was programmed using the z-Tree software (Fischbacher, 1999). Participants were undergraduate students from different disciplines at the University of Jena. After being seated at a computer terminal, participants received written instructions (see the Appendix for an English translation). In the six treatments comprising the public goods experiment, the instructions distributed at the beginning explained the rules of the public goods game only. Understanding of the rules was ensured by a control questionnaire that subjects had to answer before the experiment started. Since the payoffs attained in the public goods experiment might affect behavior in the investment game, subjects received no feedback after the public goods game.[7]

[7] On the relevance of group performance for the establishment of group identity, see Worchel et al. (2000).

In total, we ran ten sessions. Two sessions were devoted to check whether our manipulations of group identity did produce the intended effects. The remaining eight sessions were run to investigate how individual identification with a group and trust in its other members are interrelated. In all experiments amounts were denoted by ECU, where 10 ECU equal €2 in the public goods game and €5 in the investment game. The sessions with the investment game only lasted about 1 hour (with average earnings of €7.20), and the sessions with the public goods game preceding the investment game about $1\frac{1}{2}$ hours (with average earnings of €16.78).

4 Experimental results

In reporting our results, we proceed as follows. First, we examine the efficacy of the experimental manipulations. Then, we explore their impact on trust and trustworthiness. Finally, we try to identify some features of individual behavior by studying participants' choices in more depth.

4.1 Manipulation checks

To check whether randomly assigning subjects to X- or Y-type suffices to establish group membership, and whether subjects develop a sense of common fate within and/or between subgroups before playing the investment game, we recruited 62 undergraduates in two separate sessions. One session, with 30 participants, was devoted to the labeling treatments (with eight and seven 2-person groups interacting in *LAB-IN* and *LAB-OUT*, respectively), and the other, with 32 participants, to the public goods games (with four 4-person groups playing either the ϵ- or the 9ϵ-treatment).

In these sessions, following the interaction and before receiving feedback about experimental earnings, participants answered a questionnaire containing items asking for

(a) five evaluative ratings (on a scale from 1, *not at all*, to 7, *very much*) of in- and outgroup member(s) regarding how much they liked and identified with (each of) the other(s) as well as how honest, cooperative, and valuable they thought each would be;

(b) their conceptual representation ("one group" or "separate individuals" in the labeling treatments, and also "two subgroups within one larger group" or "two groups" in the public goods games) of the aggregate as well as to what extent (on a 1–7 scale) the aggregate felt like each of the representations.[8]

As a check of the common fate manipulations, (similar to Gaertner et al., 1999) participants in the public goods game were also asked to indicate the extent to which they agreed with the following statement: "If the two persons of the same type contribute to their own project, the two persons of the other type will also benefit from this".

The results from these two *ad hoc* sessions are summarized in Table 2, which displays separate indexes for in- and outgroup members, consisting of the mean of subjects' five evaluative ratings of (each of) the other(s).[9]

<u>Insert Table 2 about here</u>

Pooling all treatments, the ingroup member receives higher evaluations than the outgroup member(s) ($M_{\text{ingroup}} = 4.66$, $SD = 0.36$; $M_{\text{outgroup}} = 3.66$, $SD = 0.05$). Concerning the ϵ and 9ϵ-treatments, a 2 (treatments) $\times$ 2 (ingroup and outgroup) repeated measures multivariate analysis of variance (MANOVA) on the evaluative ratings reveals a main effect for treatment (multivariate $F(1, 60) = 3.34$, $p < 0.05$) and for group (multivariate

[8]These measures are derived from Gaertner et al. (1989).

[9]To assess the reliability of our scale and the correlation among our five items, we measured Cronbach's alpha. Pooling all sessions, we get $\alpha = 0.749$. Taking each session separately does not radically alter this result, though α is closer to the critical threshold of 0.70 in the labelling treatments than in the public goods games.

$F(1, 60) = 8.41$, $p < 0.001$), meaning that evaluations are higher in ϵ than in 9ϵ, and for ingroup members than for outgroup members.[10] A similar MANOVA for the labeling treatments shows that, even when the distinction relies only on label, differences in evaluations of ingroup and outgroup members are highly significant (multivariate $F(1, 28) = 3.53$, $p < 0.01$). Hence, though artificial in nature, these groups were temporarily relevant for the participants.

The manipulation of common fate also proved to be successful. Bias (i.e., the difference between ingroup and outgroup evaluation) significantly varies across treatments ($F(2, 227) = 5.90$, $p < 0.01$): it is lowest in the ϵ-treatment, followed in turn by the labeling treatment (with no common fate), and then by the 9ϵ-treatment. That participants differentially recognize the degree of mutual fate between subgroups is corroborated by their rating of the proposed statement, which is higher in ϵ ($M = 5.88$, $SD = 1.36$) than in 9ϵ (M=4.56, $SD = 1.63$).

However, our manipulations do not seem to affect subjects' conceptual representation of the aggregate. According to Table 2, sizable percentages of participants in ϵ and 9ϵ select the one-group and the two subgroups within one larger group representations to describe their impression of the aggregate. Even though most participants in *LAB-IN* and *LAB-OUT* choose the expected representation of the aggregate (i.e., one group and separate individuals, respectively), there is no statistically detectable difference between the two representations in either treatment ($F(3, 60) = 0.37$, $p = 0.56$ for *LAB-IN*; $F(3, 60) = 0.96$, $p = 0.34$ for *LAB-OUT*).

[10] All multivariate tests are based on Wilk's criterion.

4.2 Impact of induced group identities

To explore whether the experimentally induced group identities affect the degree to which people trust, we ran eight sessions with 32 participants each. One session was devoted to each of the treatments reported in Table 1, except for the two (pure) labeling treatments (*LAB-IN* and *LAB-OUT*) which were performed in one session. Tables 3 and 4 show investment decisions in these eight sessions. Table 3 refers to the trustor and reports absolute amounts sent (i.e., x_A). Table 4 refers to the trustee and reports the ratio between reciprocated and received amount (i.e., $x_B/3x_A$).

Insert Tables 3 and 4 about here

4.2.1 The impact of mere labeling

According to the minimal group paradigm, an arbitrary ingroup-outgroup classification should suffice to activate ingroup-favoring behavior. This is not supported by the first three rows of Tables 3 and 4. Overall average amounts sent are 4.44, 4.13, and 4.50 in *C*, *LAB-IN*, and *LAB-OUT*, respectively; the respective overall average percentages returned (for positive amounts sent) are 33%, 22%, and 24%. Kolmogorov-Smirnov tests (two-sided) comparing amounts sent and amounts returned in *C* vs. *LAB-IN* or *LAB-OUT*, and in *LAB-IN* vs. *LAB-OUT* confirm that there is no difference between treatments at the 5 percent significance level.

Result 1 *Manipulating group identity via mere labels does not change participants' attitudes to trust and trustworthiness.*

4.2.2 The impact of outcome interdependence

Will trust and reciprocity be different when the ingroup-outgroup distinction relies not only on labels but also on shared interests? Before answering this question, we shortly analyze behavior in the public goods experiment. Table 5 displays the average, the median and the standard deviation of contributions for each level of outcome interdependence.[11]

Insert Table 5 about here

The most scattered contributions (see standard deviations in Table 5) are observed in the intermediate 3ϵ-treatment whose mean and median are the lowest. According to two-sided Mann-Whitney tests, we cannot reject the null hypothesis of equal contributions for ϵ vs. 9ϵ ($p = 0.88$), while we can reject it for 3ϵ vs. ϵ ($p = 0.03$) and for 3ϵ vs. 9ϵ ($p = 0.02$). This is confirmed by a generalized linear regression (relying on a negative binomial distribution) with the amount contributed as dependent variable and treatment dummies as independent variables. Among the considered explanatory variables, only the coefficient of the 3ϵ-dummy is significantly negative ($\beta = -0.347$, $p = 0.015$). Thus, average contributions are lower in the 3ϵ-treatment than in the ϵ- and 9ϵ-treatments.

Let us now examine how differences in outcome interdependence translate into trust behavior. Starting with the trustors' behavior, Table 3 shows that the non-monotonicity observed in the public goods experiment extends to the investment game: the 3ϵ-treatments (3ϵ-IN and 3ϵ-OUT) always yield the lowest average x_A-amounts, whereas the standard deviations of the investment decisions do not exhibit a clear-cut pattern. Kolmogorov-Smirnov tests (two-sided) reveal, however, that differences in intra/intergroup mutual fate do not affect the level of amounts sent in statistically significant ways ($p > 0.5$ for any comparison of treatments). Further non-parametric tests comparing trustors' decisions in the six experimental treatments with

[11] We can pool the data from the IN and OUT treatments and, hence, obtain 64 independent observations for each value of $\alpha - \beta$.

those in the control treatment indicate no significant difference in amounts sent ($p > 0.82$ in any comparison).

Result 2 *Trustors do not react to differences in outcome interdependence: in no treatment they send significantly different amounts.*

Thus, according to our data, manipulations of group identity via outcome interdependence do not influence participants' level of trust. This suggests that participants in our experiment tend to be social, rather than particularized, trusters. It has been claimed that people who believe that most other people in their society can be trusted, regardless of the extent of group boundaries, are also more inclined to give to charity or to voluntarily contribute (cf., e.g., Rothstein and Uslaner, 2005). In line with this argument, we find that amounts sent are significantly and positively correlated with the trustors' contributions in the public goods experiment. The Pearson correlation coefficient between contribution and investment decisions is 0.44 ($p < 0.01$). Such a positive relation between trust and willingness to contribute is corroborated by the results of a generalized linear regression with trustors' sending decision as dependent variable, and amount contributed to the public good and treatment dummies as independent variables.[12] There is, in fact, a highly significant positive effect of contributions ($\beta = 0.091$, $p < 0.01$), whereas the various dummy coefficients are not significant. Controlling for possible interaction effects does not alter the results.

Turning to the trustees' reciprocating behavior (see Table 4), it follows the same non-monotonic pattern observed for contribution decisions and amounts sent: the intermediate 3ϵ-treatments trigger the most extreme behavior, although the effects differ between IN and OUT (the amounts

[12]Due to censored observations, we assume a negative binomial distribution so as to model over-dispersion.

returned are the highest in 3ϵ-IN and the lowest in 3ϵ-OUT). Kolmogorov-Smirnov tests (two-sided) confirm that our manipulations of social identity affect significantly trustees' choices only when $\alpha - \beta = 3\epsilon$ ($p = 0.05$ for 3ϵ-IN vs. 3ϵ-OUT; $p > 0.4$ for any other comparison). Therefore, the ingroup-favoring behavior advocated by the minimal group paradigm is detected only for trustees, and is restricted to the intermediate 3ϵ-treatments.

Result 3 *Trustees reciprocate significantly less in the OUT-treatment than in the IN-treatment only when $\alpha - \beta = 3\epsilon$. In case of extreme $\alpha - \beta$ values, group identity manipulations do not trigger significantly different reciprocity.*

The correlation between amounts contributed to the public good and amounts returned in the investment game is positive but not significant at the 10 percent level. This is confirmed by the results of a generalized linear model (assuming a negative binomial distribution) regressing the amount returned by trustees (i.e., x_B) on the amount sent by trustors, the contribution to the public good by trustees, and treatment dummies. Whereas the amount received has a significantly positive effect on the amount returned ($\beta = 0.191$, $p < 0.01$), implying that reciprocity is important, former contribution decisions have no significant impact on x_B. Moreover, among all treatment dummies, only the coefficient of the interaction term between matching procedures (IN and OUT-treatments) and 3ϵ is significantly positive, implying (in line with Result 3) that trustees tend to return more when matched with an ingroup, rather than an outgroup, member only in the 3ϵ-treatment.

4.3 A closer look at the individual data

We first focus on trustors, and rely on the following classification: *non-trust* (or *opportunism*) if $x_A = 0$, *weak trust* if $0 < x_A \leq 5$, *strong trust* if

$5 < x_A < 10$, and *full trust* if $x_A = 10$. Table 6 reports the proportion of trustors' choices in each treatment.

Insert Table 6 about here

In line with previous findings, the distribution of choices is very similar in all treatments with weak trust being always dominant. Except for ϵ-OUT, where we observe a higher proportion of strong trustors, $x_A \leq 5$ is the most prevalent behavior. The percentage of trustors sending exactly half of their endowment is the mode in any treatment.

Examining the relationship between contribution and sending decisions of the various types, we find that contribution levels of weak and strong trustors are not statistically significantly different ($p = 0.246$, Kolmogorov-Smirnov test), but contribution levels of the other types differ at the 5 percent significance level.

Turning to trustees, based on the return ratio $r = x_B/3x_A$ (for $x_A > 0$), we classify their choices as follows: *non-reciprocity* (or *opportunism*) if $r = 0$, *half-exploitation* if $0 < r < 1/3$ (or $x_B < x_A$), *weak reciprocity* if $r = 1/3$ (or $x_B = x_A$), and *strong reciprocity* if $r > 1/3$ (or $x_B > x_A$). Table 7 depicts the proportion of trustees' choices in each treatment.

Insert Table 7 about here

Quite generally, choices tend to concentrate on extreme values. In no treatment trustees send back the amount received (weak reciprocity is never detected).[13] The percentage of strong reciprocal choices is 58.33%, 53.33%, and 69.23% in C, ϵ-IN, and 3ϵ-IN, respectively. By contrast, in all the other treatments, this percentage is at the most 50%, and reaches its minimum

[13]This finding stands against the $x_A = x_B$-modal behavior observed by Glaeser et al. (2000). However, Glaeser et al.'s experiment differs from ours in that trustees receive no endowment and the multiplier equals two.

(27.27%) in 3ϵ-OUT, where the overwhelming proportion of choices (45.45%) is opportunistic. This confirms that only in the intermediate 3ϵ-treatment trustworthiness favors ingroup members.

Do the four different types of trustees contribute differently to the public good (in the six respective treatments)? A series of Kolmogorov-Smirnov tests (two-sided) comparing the distributions of contributions by the various types reveals no statistically significant differences at the 10 percent level.

5 Discussion

We have experimentally investigated how individual identification with a group may be related to trust in its other members. Group identity was manipulated through two channels: by creating a minimal ingroup-outgroup distinction (merely based on labels), and via positive outcome interdependence established in a prior public goods game.

The results from the *ad hoc* sessions performed to check our manipulations' effectiveness indicate that participants did feel greater identity with ingroup than with outgroup members, and that common fate reduced intergroup bias. However, unlike earlier experiments employing a minimal group paradigm and dealing with simple allocation tasks, we found no significant effects of induced group identity on trust decisions. Only for the intermediate level of outcome interdependence, trustees exhibited the predicted ingroup-favoring behavior. Furthermore, individual preferences towards cooperation appeared to be quantitatively more important than group membership. We observed, indeed, a significant positive correlation between amounts sent in the investment game and contribution decisions in the public goods game. The same qualitative conclusions held for the amounts reciprocated, although the effects were insignificant. These results are in line with those by Gächter et al. (2004), who found contributions in public goods experiments to be significantly correlated with three trust indicators.

Should we be disappointed because we did not find any convincing evidence of ingroup-favoring behavior? In our view, the evidence that people have faith in those who are different from themselves is good, rather than bad, news. It means that not every weak possibility of partitioning a society leads to intra-societal conflicts, and that social trust is behaviorally relevant.

Many researchers in all social sciences have by now acknowledged that social trust influences a wide range of economic and political phenomena. Without trust, no market could function (Arrow, 1974), which explains, for example, why Fukuyama (1995) concludes that social capital or trust is as important as physical capital in establishing large-scale business organizations necessary for economic growth and development. It has also been argued that social trustors tend to be more optimistic about their own ability to influence their life chances, and also happier with their present life (cf., Delhey and Newton, 2003). Hence, at both the societal and the individual levels, many normatively desirable features seem connected to social trust. While according to the economic literature trust can be achieved in a roundabout way through institutional and other devices incentivizing trustworthiness (see James, 2002, for a survey of solutions to the trust problems proposed by economists), our experiment shows that, even without such incentives for trustworthiness and in the presence of social boundaries, people have a predisposition to *just* trust each other regardless of subgroup affiliation.

A puzzling result of our experiment is the non-monotonic interdependence of contribution and trust behavior since the intermediate 3ϵ-treatments always triggered the most extreme behavior. An insight into this observation comes from previous evidence (Breakwell, 1978), suggesting that individuals who are characterized by a degree of group membership in-between two extremes may perceive their identity as threatened with respect to the others and, consequently, focus on positive traits of their group. From a methodological point of view, such non-monotonicity illustrates that parameters should be varied more systematically when trying to understand what drives experimental behavior.

Appendix: Experimental instructions

This appendix reports the instructions (originally in German) we used for the labeling and ϵ-treatments. The instructions for the other treatments were adapted accordingly and are available upon request.

General instructions (common to all treatments)

Welcome and thanks for participating in this experiment. Please read the following instructions carefully. From now on any communication with other participants is forbidden. If you have any questions or concerns, please raise your hand. We will answer your questions individually. It is very important that you follow this rule, otherwise we will exclude you from the experiment and from all payments.

The experiment allows you to earn money. Your experimental income will be calculated in ECU (Experimental Currency Unit), where 1 ECU = [*In the ϵ-treatments:* €0.20] [*In the LAB-treatments:* €0.50]. At the end of the experiment, the ECU-income you have earned will be converted to Euro and paid to you in cash.

Instructions for the public goods game (ϵ-treatments)

In this experiment, you are randomly matched with three other persons, whose identity we will not reveal to you at any time. Two of you will be of type X and two of you of type Y. You will learn your type before the experiment starts.

Each participant − regardless of his/her type − receives an endowment of 10 ECU, and must face **only once** the following choice situation.

- X-types must decide how many ECU they want to contribute to project X.
- Y-types must decide how many ECU they want to contribute to project Y.

Whatever is not contributed is kept for oneself. The sum of all contributions to X is called X-amount. The sum of all contributions to Y is called Y-amount. Your earnings is the sum of

(1) the "ECU you keep"; i.e.: 10 ECU − your contribution;

(2) your "income from the projects":

Your earnings = (**10**−**your contribution**) + **your income from the projects**.

Your "**income from the projects**" is determined as follows:

$$\textit{Income from the projects for } X\textit{-types} = \boxed{\begin{array}{c} \mathbf{0.55} \times [\mathbf{X\text{-}amount}] \\ + \\ \mathbf{0.45} \times [\mathbf{Y\text{-}amount}] \end{array}}$$

$$\textit{Income from the projects for } Y\textit{-types} = \boxed{\begin{array}{c} \mathbf{0.45} \times [\mathbf{X\text{-}amount}] \\ + \\ \mathbf{0.55} \times [\mathbf{Y\text{-}amount}] \end{array}}$$

Thus both X-types receive the same income from the projects. If, for example, the X-amount is 20 ECU and the Y-amount is 10 ECU, X-types receive $(0.55 \times 20) + (0.45 \times 10) = 11 + 4.5 = 15.5$ ECU. Likewise, the income from the projects is the same for all Y-types; i.e., both Y-types receive the same income from the projects. If, for example, the X-amount is 1 ECU and the Y-amount is 10 ECU, Y-types receive $(0.45 \times 1) + (0.55 \times 10) = 0.45 + 5.5 = 5.95$ ECU.

Note that the contribution by an X-type increases only the X-amount. Likewise, the contribution by a Y-type increases only the Y-amount. If you are a participant of type X and contribute, for instance, 1 ECU, this increases the X-amount by 1 ECU and leaves the Y-amount unchanged. As a consequence, your income as well as the income of the other X-type increase by 0.55 ECU, and the income of the two Y-types increases by 0.45 ECU.

If you are a participant of type Y and contribute, for instance, 1 ECU, this increases the Y-amount by 1 ECU and leaves the X-amount unchanged. As a consequence, your income as well as the income of the other Y-type increase by 0.55 ECU, and the income of the two X-types increases by 0.45 ECU.

Each ECU that you keep for yourself raises "ECU you keep". Thus, each ECU that you keep yields money for YOU ALONE. The others do not receive anything for the

ECU that you keep.

You will receive information about the number of ECU contributed by the others and your earnings at the end of today session.

Before the experiment starts, you will have to answer some control questions to verify your understanding of the experiment.

Please remain quiet until the experiment starts and switch off your mobile phone. If you have any questions, please raise your hand now.

Instructions for the investment game

[*Participants in the ϵ-treatments read:* Please read the instructions carefully. If you have any questions or concerns, please raise your hand. As before, your experimental income will be calculated in ECU but now 1 ECU = €0.50.

In this experiment, participants are randomly divided into pairs composed by persons who were of the **same type** in experiment I. That is, if you were an X-type in the previous experiment, you will be paired with the other X-type; if you were a Y-type in the previous experiment, you will be paired with the other Y-type.

In ϵ-OUT this paragraph was replaced by: In this experiment, participants are randomly divided into pairs composed by persons who were of **different types** in experiment I. That is, if you were an X-type in the previous experiment, you will be paired with one of the two Y-types; if you were a Y-type in the previous experiment, you will be paired with one of the two X-types.]

[*Participants in the LAB-treatments read:* The participants in the experiment can be of two types: either X or Y. **Please open the attached white envelope to learn about your type**. In the experiment, you will be interacting with a person of your **SAME type**. That is, if you are an X-type, you will be paired with an X-type; if you are a Y-type, you will be paired with a Y-type.

In LAB-OUT the last two sentences were replaced by: In the experiment, you will be interacting with a person of a **DIFFERENT type**. That is, if you are an X-type, you will be paired with a Y-type; if you are a Y-type, you will be paired with an X-type.]

The two members of a pair will be randomly assigned one of two roles: either A or B. Each A-person and each B-person will receive an endowment of 10 ECU, and must take **only one decision**.

- *A-participants' task*

As an A-person, you must decide **how much of your endowment you want to give to B**. You can choose only integer amounts; i.e., 0, 1, 2, $\cdots$, 9 or 10 ECU. For any amount you decide to give to B, what B actually receives is the triple. If you give, for instance, 1 ECU, B receives 3 ECU. If you give 3 ECU, (s)he receives 9 ECU. If you give 5 ECU, (s)he receives 15 ECU.

- *B participants' task*

As a B-person, you are informed about the number of ECU you receive from A, which is three times what A has given to you. You must then decide **how much of the received amount you want to give back to A**. If A has given you, for example, 5 ECU and thus you have received 15 ECU, you must decide how many of these 15 ECU you want to return to A. Only integer amounts can be sent back.

The earnings of A and B are given by:

A's earnings	=	**10**	−	[**what (s)he gives to B**]
			+	[**what (s)he receives from B**]
B's earnings	=	**10**	+	[**3 × what (s)he has been given by A**]
			−	[**what (s)he returns to A**]

Your role (A or B) will be told to you before the experiment starts.

After all participants have made their choice, you will receive information about your earnings as well as about the number of ECU you received from B if you are an A-person.

Please remain quiet until the experiment starts. If you have any questions, please raise your hand now.

References

Arrow, K., 1974. The Limits of Organizations. Oxford University Press, New York.

Berg, J., Dickhaut, J., McCabe, K., 1995. Trust, reciprocity, and social history. Games and Economic Behavior 10, 122–142.

Breakwell, G., 1978. Some effects of marginal social identity. In: Tajfel, H. (Ed.), Differentiation between Social Groups. Academic Press, London.

Brewer, M.B., 1979. In-group bias in the minimal intergroup situation: a cognitive-motivational analysis. Psychological Bulletin 86, 307–324.

Brewer, M.B., 1981. Ethnocentrism and its role in interpersonal trust. In: Brewer, M.B., Collins, B. (Eds.), Scientific Inquiry in the Social Science. Jossey-Bass, San Francisco, CA.

Brown, R., 2000. Social identity theory: past achievements, current problems and future challenges. European Journal of Social Psychology 30(6), 745–778.

Buchan, N., Croson, R., 2004. The boundaries of trust: own and other's actions in the US and China. Journal of Economic Behavior & Organization 55, 485–404.

Buchan, N., Johnson, E.J., Croson, R., 2006. Let's get personal: an international experiment of the influence of communication, culture, and social distance on other regarding preferences. Journal of Economic Behavior & Organization 60(3), 373–398.

Cox, J.C., 2004. How to identify trust and reciprocity. Games and Economic Behavior 46, 260–281.

Dayton-Johnson, J., 2003. Knitted warmth: the simple analytics of social cohesion. Journal of Socio-Economics 32, 623–645.

Delhey, J., Newton, K., 2003. Who trusts? The origins of social trust in seven societies. European Societies 5(2), 93–137.

Fischbacher, U., 1999. Zurich toolbox for readymade economic experiments. Working Paper No. 21, Institute for Empirical Research in Economics, University of Zurich.

Frey, B.S., Bohnet, I., 1997. Identification in democratic society. The Journal of Socio-Economics 26, 25–38.

Fukuyama, F., 1995. Trust: The Social Virtues and the Creation of Prosperity. The Free Press, New York.

Gächter, S., Herrmann, B., Thöni, C. 2004. Trust, voluntary cooperation, and socio-economic background: survey and experimental evidence. Journal of Economic Behavior and Organization 55, 25–30.

Gaertner, S., Dovidio, J., Rust, M., Nier, J., Banker, B., Ward, C., Mottola, G., Houlette, M., 1999. Reducing intergroup bias: elements of intergroup cooperation. Journal of Personality and Social Psychology 76, 388–402.

Gaertner, S., Mann, J., Murrell, A., Dovidio, J., 1989. Reducing intergroup bias: the benefits of recategorization. Journal of Personality and Social Psychology 57, 239–249.

Glaeser, E.L., Laibson, D., Scheinkman, J.A., Soutter, C.L., 2000. Measuring trust. Quarterly Journal of Economics 115(3), 505–531.

Güth, W., Levati, M.V., Sausgruber, R., 2005. Tax morale and (de)centralization: an experimental study. Public Choice 125, 171—188.

Hornsey, M.J., van Leeuwen, E., Van Santen, W., 2003. Dragging down and dragging up: how relative group status affects responses to common fate. Group Dynamics: Theory, Research, & Practice 7(4), 275–288.

James Jr., H.S., 2002. The trust paradox: a survey of economic inquiries into the nature of trust and trustworthiness. Journal of Economic Behavior & Organization 47, 291–307.

Jin, N., Yamagishi, T., 1997. Group heuristics in social dilemma. Japanese Journal of Social Psychology 12, 190–198.

Kramer, R.M., 1991. Intergroup relations and organizational dilemmas: the role of categorization processes. In: Cummings, L.L., Staw, B.M. (Eds.), Research in Organizational Behavior. JAI Press, Greenwich.

Kramer, R.M., Brewer, M.B., 1984. Effects of group identity on resource use in a simulated commons dilemma. Journal of Personality and Social Psychology 46, 1044–1057.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R., 1997. Trust in large organization. American Economic Review Papers and Proceedings 87(2), 333–338.

Orbell, J.M., van de Kragt, A., Dawes, R.M., 1988. Explaining discussion-induced cooperation. Journal of Personality and Social Psychology 54(5), 811–819.

Ortmann, A., Fitzgerald, J., Boeing, C., 2000. Trust, reciprocity, and social history: a re-examination, Experimental Economics 3, 81–100.

Putnam, R., 1993. Making Democracy Work: Civic Tradition in Modern Italy. Princeton University Press, Princeton, NJ.

Putnam, R., 2000. Bowling Alone: The Collapse and Revival of American Community. Simon & Schuster, New York.

Rosenbaum, M. E., Moore, D. L., Cotton, J. L., Cook, M. S., Hieser, R. A., Shovar, M. N., Gray, M. J., 1980. Group productivity and process: pure and mixed reward structures and task interdependence. Journal of Personality and Social Psychology 39, 626–642.

Rothstein, B., 2005. Social Traps and the Problem of Trust. Cambridge University Press, Cambridge.

Rothstein, B., Uslaner, E. M., 2005. All for all: equality, corruption, and social trust. World Politics 58, 41–72.

Schopler, J., Insko, C.A., 1992. The discontinuity effect in interpersonal and intergroup relations: generality and mediation. European Review of Social Psychology 3, 121–151.

Singleton, S., Taylor, M., 1992. Common property, collection action and community. Journal of Theoretical Politics 4, 309–324.

Tajfel, H., 1970. Experiments in intergroup discrimination. Scientific American 223(5), 96–102.

Tajfel, H., Billig, M., Bundy, R.P., Flament, C., 1971. Social categorization and intergroup behavior. European Journal of Social Psychology 1, 149–177.

Uslaner, E.M. 1999. Democracy and social capital. In: Warren, M. (Ed.), Democracy and Trust. Cambridge University Press, Cambridge.

Uslaner, E.M. 2000. Producing and consuming trust. Political Science Quarterly 115(4), 569–590.

Worchel, S., Iuzzini, J, Countant, D., Ivaldi, M., 2000. A multidimensional model of identity: relating individual and group identities to intergroup behavior. In: Capozza, D., Brown, R. (Eds.), Social Identity Processes. Sage Publications, Thousand Oacks, CA.

Table 1: Summary of experimental design

	Public Goods Game	Investment Game
Treatment	$\alpha - \beta$	*Pairs*
C (Control)	–	random
LAB-IN	–	same label
LAB-OUT	–	different label
ϵ-*IN*	0.05	same label/subgroup
3ϵ-*IN*	0.30	same label/subgroup
9ϵ-*IN*	0.90	same label/subgroup
ϵ-*OUT*	0.05	different label/subgroup
3ϵ-*OUT*	0.30	different label/subgroup
9ϵ-*OUT*	0.90	different label/subgroup

Table 2: Effects of labeling and intra/intergroup shared interests on measures of ingroup identification and representation of the aggregate

Measure	Treatments			
	LAB-IN ($\nu^a = 16$)	*LAB-OUT* ($\nu = 14$)	ϵ ($\nu = 16$)	9ϵ ($\nu = 16$)
Rating of group members				
Ingroup_{index}	4.57		4.36	5.06
Outgroup_{index}		3.61	3.71	3.66
Bias	0.96		0.65	1.40
Representations of the aggregate				
Percentage				
Separate individuals	37.5%	57.0%	12.5%	18.75%
One group	62.5%	43.0%	50.0%	37.50%
Two subgroups within one group			37.5%	31.25%
Two groups			0.0%	12.50%
Mean rating				
Separate individuals	3.81	4.64	2.56	3.50
One group	4.25	3.79	5.13	4.56
Two subgroups within one group			4.56	5.19
Two groups			2.88	3.38

[a] ν denotes the number of observations.

Table 3: Average absolute amount sent by trustors in each treatment

Treatment	ν^a	Mean	Median	Std deviation
Control	16	4.44	4.50	3.72
LAB-IN	8	4.13	4.50	2.03
LAB-OUT	8	4.50	5.00	3.50
ϵ-*IN*	16	5.06	4.50	2.95
3ϵ-*IN*	16	3.75	4.00	2.96
9ϵ-*IN*	16	4.31	4.00	2.94
ϵ-*OUT*	16	5.13	5.50	3.44
3ϵ-*OUT*	16	3.38	2.50	3.32
9ϵ-*OUT*	16	4.69	4.00	2.98

[a] ν denotes the number of observations.

Table 4: Average relative amount returned by trustees in each treatment

Treatment	ν^a	Mean	Median	Std deviation
Control	12	0.33	0.37	0.28
LAB-IN	8	0.22	0.13	0.27
LAB-OUT	6	0.24	0.23	0.26
ϵ-*IN*	15	0.26	0.33	0.24
3ϵ-*IN*	13	0.33	0.33	0.28
9ϵ-*IN*	14	0.20	0.09	0.22
ϵ-*OUT*	14	0.24	0.29	0.21
3ϵ-*OUT*	11	0.15	0.06	0.17
9ϵ-*OUT*	14	0.21	0.12	0.22

[a] ν has the same interpretation as in Table 3.

Table 5: Average contributions to the public good

Treatment	Mean	Median	Std deviation
ϵ	6.41	7.50	3.69
3ϵ	4.87	5.00	3.93
9ϵ	6.34	6.00	3.53

Table 6: Trustors' percentage of choices in accordance with each type (I: non-trust, II: weak trust, III: strong trust, IV: full trust)

Treatment	I	II	III	IV
Control	4	6	3	3
LAB-IN	0	6	2	0
LAB-OUT	2	4	1	1
ϵ-*IN*	1	10	2	3
3ϵ-*IN*	3	9	3	1
9ϵ-*IN*	2	10	2	2
ϵ-*OUT*	2	6	5	3
3ϵ-*OUT*	5	8	1	2
9ϵ-*OUT*	2	9	4	1

Table 7: Trustees' percentage of choices in accordance with each type (I: opportunism, II: half-exploitation, III: weak reciprocity, IV: strong reciprocity)

Treatment	I	II	III	IV
Control	3	2	0	7
LAB-IN	4	1	0	3
LAB-OUT	3	0	0	3
ϵ-*IN*	5	2	0	8
3ϵ-*IN*	4	0	0	9
9ϵ-*IN*	5	4	0	5
ϵ-*OUT*	5	2	0	7
3ϵ-*OUT*	5	3	0	3
9ϵ-*OUT*	6	2	0	6