

Identity, Cooperation, and Punishment*

Kendra N. McLeish[†] Robert J. Oxoby[‡]

March 2006

Abstract

Among economists, there is increased recognition of the role individuals' identities play in decision-making. In this paper, we conduct laboratory experiments in which we explore the motivations for and the effects of group identity. We find that negative out-group opinion (acting as an inter-group identity threat) can motivate in-group/out-group effects in a simple bargaining context. Further, our results suggest that disparagement of group norms by members of the in-group (acting as an intra-group identity threat) increases the use of costly punishment within the in-group.

KEYWORDS: identity, fairness, reciprocity, experiments

JEL CODES: C9, D1, M5

***Preliminary Version - Comments Welcomed.** We thank George Akerlof, Alexander Haslam, Robert Putnam, and participants at the CIAR workshop on Social Interaction, Identity, and Well-Being (March 2006) for helpful comments. We thank Suzanne Lamoureaux, Beata Komierowski, Alexander Smith, and Teresa Wong for research assistance. Financial support from the Institute for Advanced Policy Research and the Social Science and Humanities Research Council of Canada is acknowledged.

[†]Department of Economics, University of Calgary, 2500 University Drive NW, Calgary AB Canada T2N 1N4; knmcleis@ucalgary.ca.

[‡]Corresponding author: Department of Economics, University of Calgary, 2500 University Drive NW, Calgary AB Canada T2N 1N4; tel (403) 220-2586; fax (403) 282-5262; oxoby@ucalgary.ca.

1 Introduction

Individuals often perceive others in a biased manner, frequently forming perceptions based on the social groups for which an individual is a member. Individuals identify with one another through these groups, impacting expectations of others' behavior as well as the behavioral norms governing interactions. Thus, a given decision environment may yield different outcomes depending on the group membership of those involved. By permitting individuals to better anticipate the behaviors of others, the presence of identity in decision making can also reduce transaction costs and facilitate interactions. Indeed, a significant research has demonstrated the manners in which group membership and identification affects individual behavior and perception. For example, group identity has been shown to influence attitude formation (Mackie et al., 1990), cooperation (Wit and Wilke, 1992), reciprocity (Stroebe et al., 2005), and negotiations (Kramer et al., 1993). On the other hand, identity, particularly working through unfavorable out-group biases may result in polarization (Mackie, 1986) and the differential treatment of out-group members with whom one does not identify as strongly (Bernard et al., 2006; Durlauf, 1999; Gerber, 1998; Wann and Grieve, 2005).¹

Social scientists often look at intergroup relations through the lens of social identity (Tajfel et al., 1971; Tajfel and Turner, 1986). Identity or one's self-concept is determined, in part, by one's membership in and perceived similarity with various social groups. This identification as a member of a social unit or group transforms an individual's self-concept, promoting the internalization of norms and attitudes which in turn affect behavior in subsequent social interactions (see Haslam, 2001). Once individuals identify with certain groups, individuals seek to protect this self-concept by looking for evidence or acting in a manner which preserves the positive aspects of this identity. That is, individuals seek to protect their self-concept. Threats to one's own

¹This type of effect is closely related to the idea of statistical discrimination and negative stereotyping. See Lundberg and Startz (1983) and Coate and Loury (1993).

identity or the identity of a group (e.g. changes to the status of the group) often serve as stressors that motivate inter-group competition (Mackie, 1986), group disassociation (Matheson and Cole, 2004), discrimination towards the out-group (Wann and Grieve, 2005), or amplified identification with the in-group (Scheepers and Ellemers, 2005; Wann and Grieve, 2005).² In addition to behavior being shaped by threats to group identity from outside the group (inter-group identity threat), Akerlof and Kranton (2000) emphasizes how behavior is shaped by threats to identity from inside the group (intra-group identity threat generated by the violation of in-group norms). Formalizing this idea in economic terms, Akerlof and Kranton (2000) posit that group identity imposes a payoff interdependency between individuals: If a member of the group acts in a way that threatens the identity of the group by violating group norms, group members may act to try to protect that identity.

In this paper, we conduct experiments to test the manner in which identity and identity threat influence decision making in a series of simple two-person interactions. We hypothesize that the presence of identity should result in a favorable in-group bias which manifests itself in greater cooperation with members of one's in-group. Moreover, identity threat in the form of the violation of in-group norms (i.e. intra-group identity threat) will result in behaviors which are intended to reaffirm one's identity.

In our experiments, we have groups of participants engage in an identity-building task, introduce inter-group identity threat through a status manipulation, and then test patterns of cooperation and punishment amongst in- and out-group pairs. This design permits us to analyze not only the raw effect of identity on decision making (i.e. how interacting with an in-group member differs from interacting with an out-group member) but also allows us to focus on the importance of inter-group identity

²The effects of identity are not restricted to behaviors, attitudes or perceptions: Matheson and Cole (2004) and Scheepers and Ellemers (2005) document how group status manipulations can motivate physiological threat responses in subjects.

threat in decision making. In our context, inter-group identity threat is motivated by perceptions of one's in-group by out-group members.

We argue that these types of perceptions or opinions on the part of out-group members increase one's attention to their own identity and results in greater identity-based decision making (e.g. greater in-group favoritism as documented by Stroebe et al., 2005; Wann and Grieve, 2005). Responses to inter-group identity threat are prevalent among high status groups who must defend their positions as well as among low-status groups (Scheepers and Ellemers, 2005). Relatedly, research has shown that inter-group identity threat may cause group disassociation when a group (particularly, a low-status group) believes its position is justifiable (Branscombe et al., 2002). Conversely, inter-group identity threat is amplified (leading to greater in-group identification and out-group hostility) when a group sees its position as unmerited. For example, Branscombe et al. (1999) argue that women are more likely to accept discrimination relative to men on the basis of physical strength, while African Americans are less likely to accept discrimination relative to Whites on the basis of ability.³ In fact, Branscombe et al. (1999) show that African American group identification increases as the validity of the status comparison with White Americans becomes more questionable.

Our results provide two important insights into the role of identity threat in economic and social interactions. First, when identity is sufficiently motivated through out-group opinion on status characteristics (i.e. inter-group identity threat) we observe individuals engaging in greater cooperation with those in their in-group. Specifically, the more that out-group opinion ranked the in-group lower than participants in the in-group ranked themselves, the more participants cooperated with their in-group. Conversely, the more that out-group opinion ranked the in-group higher than

³This type of inter-group identity threat is related to the idea of stereotype threat discussed by Steele and Aronson (1995).

participants in the in-group ranked themselves, the less participants cooperated with their in-group. Secondly, we observe greater negative reciprocity among individuals in the in-group when the group's norm of behavior is violated (i.e. intra-group identity threat). This latter result conforms with the instances of "acting out" discussed in Akerlof and Kranton (2000, 2005) in which individuals respond negatively to behaviors which challenge their group identity.

The effects of identity in decision making can be multifaceted. Akerlof and Kranton (2000, 2005) offer many examples of behaviors that are largely influenced by the creation and protection of group identity. As examples of the effects of identity, the authors cite the isolation (or harassment) of women working in traditionally male fields and the hostility towards women's rights supporters from fellow females. These behaviors (isolation, hostility) are brought on by identity threat emerging from either another group (e.g. women entering a male field) or across groups (e.g. women's rights supporters violating the norms of other women). Other examples of antisocial behavior resulting from threat to one's group identity (specifically, gender identity) include discrimination and being ostracized as a result of violating in-group norms.⁴ Similarly, Devine et al. (1999) argues that the social isolation faced by people with HIV and AIDS is, in part, a result of hostility towards homosexuals, seen as being deviants.

In terms of broader economic issues, understanding identity can have important implications for labor market policies (Akerlof and Kranton, 2000, 2005; Berg and Lein, 2002), the formation of human capital (Akerlof and Kranton, 2002), understanding poverty and exclusion (Bernard et al., 2006; Hoff and Pandey, 2006; Oxoby, 2004, 2005), and welfare and redistributive programs (Alesina et al., 2001; Lee and

⁴Both Badgett (1995) and Berg and Lein (2002) identify a negative wage effect of homosexuality for men (ranging from 22% to 27%). This research seemingly raises a paradox as homosexual women do not earn less (and by some accounts earn more) than heterosexual women. A possible explanation lies in the application of identity threat wherein men engaging in hiring engage in identity related discrimination solely against other males.

Roemer, 2004). Similarly, identification may be an important aspect in the formation of social capital and achieving economic growth (Glaeser et al., 2002; Putnam, 1993; Zak and Knack, 2001).⁵

The remainder of our paper is organized as follows: Section 2 reviews some of the psychological literature on the motivation of identity within and between groups and presents some of the previous experimental literature on identity in economics. In section 3 we present our experimental design and a simple model to motivate our hypotheses. Section 4 presents our results and section 5 concludes.

2 Motivating Identity

Within the social sciences, a large literature has focused on uncovering the contributors to and effects of group identity. Turner (1982) presents three factors considered fundamental for the motivation of identification among individuals (both in-group and out-group) :

1. individuals must be perceived or recognized as members of different groups;
2. individuals must learn the different norms, attitudes, or behaviors that distinguish groups;
3. individuals must engage in self-stereotyping in which they attribute the characteristics of the in-group to themselves.

In terms of the effects of identity, in-group identification yields greater perceptions of similarity, greater cooperation, and greater conformity of behaviors and attitudes within the group (Nass et al., 1996; Tajfel et al., 1971; Tajfel and Turner, 1986).

⁵Identity reflects an additional margin on which individuals and others in a heterogeneous population interact. As such, identity may be an important aspect in understanding how heterogeneity in a population affects interactions and the provision of public goods (see Alesina et al., 1999; Alesina and Ferrara, 2000).

Moreover, the effective motivation of identity appears to yield polarization of group attitudes, thereby enhancing factors 1 and 2 (in order to yield more defined separation among groups; Mackie, 1986; Mackie et al., 1990; Wilder and Shapiro, 1984).

Experimentally, the manipulation of minimal cues is often sufficient to motivate identification between individuals. The essential aspect of such cue manipulations is factor 1: the recognition of different groups (Nass et al., 1996; Mackie, 1986). This recognition appears to be strongly reinforced in environments in which there is an “interdependence of fate” within groups. For example, Rabbie and Horowitz (1969) use a simple chance at a win or loss (i.e. a coin toss) to motivate identity within groups. They find that this cue (winning or losing) is sufficient to yield substantial differences in the ways individuals perceive (i.e. form first impressions regarding) in-group and out-group members. Similarly, Wilder (1990) manipulated social identity by having participants wear badges and congregate in labeled rooms which indicated their group membership. This seemingly innocuous manipulation resulted in participants being more influenced by messages from in-group members (even when these messages were identical in all dimensions except the originating group). Mackie (1986) promoted intergroup competition, thereby motivating identity among the groups through differentiation (factor 1) and emphasizing in-group/out-group norms based on this competition (factor 2). Results indicate an identity effect regarding polarization and norm adherence: When participants listened to identical taped discussions from perceived group members, participants in the intergroup competition treatment perceived group norms to be more extreme and showed greater attitude polarization than participants in the control group.

While these studies have focused on the aspects of interactions within a group that motivate identity, other research has explored how the behaviors or beliefs of an out-

group affect in-group identity and in-group cohesion.⁶ The focus of this research has been on “inter-group identity threat” and emphasizes the import of how members of a group perceive themselves relative to other groups (factor 3). In this line of research, the focus is on how the behaviors and beliefs of one group serve to strengthen the identification within another group. In a sense, the beliefs of one group permit relative comparisons by other groups, yielding greater group recognition and in-group norm adherence. Thus, factor 3 reinforces factors 1 and 2. For example, such threats emerge in contexts in which a high status group observes gains in the circumstances of lower-status groups (Scheepers and Ellemers, 2005). These gains by another group threaten the position of the higher status group, motivating a greater sense of identity within the high status group.

Moreover, other groups’ beliefs (particularly unfavorable opinions) about one’s own group motivate stronger group recognition and self-stereotyping (factors 1 and 3). Specifically, the manner in which group biases are framed affect identification. For example, Brown and Ross (1982) finds that negative out-group opinion (of the in-group) result in greater in-group cooperation and larger in-group effects (i.e. those described by Nass et al., 1996). Branscombe et al. (1999, 2002) find that group identification is strongest when individuals of one group question the way they have been judged by another group or believe that they are treated unfairly relative to another group. The central aspect of these studies has been a focus on the validity with which in-group members consider their treatment by out-group members: when in-group members believe that they are illegitimately viewed in a negative light or treated unfairly by out-group members a stronger sense of in-group identification emerges.

⁶How the characteristics of an out-group can effect cohesion and cooperation of an in-group are discussed in Oxoby (2004, 2005).

In economic experiments with salient incentives, building group identity appears to be more difficult than research in other social sciences suggests (e.g. Eckel and Grossman, 2005). Using a manipulation similar to that in Rabbie and Horowitz (1969), Wit and Wilke (1992) presented either individuals or groups with a 50% chance of receiving tokens they had accumulated in a previous game.⁷ Results suggest that individuals only contributed slightly more in the games under the “group chance” treatment (68%) than under the “individual chance” treatment (57%). In a similar environment more germane to our experiment, Stroebe et al. (2005) find marginal support for an increase in reciprocity when there was group outcome dependencies motivating identity.

These results suggest the need for stronger identity manipulations in the face of salient incentives. To this end, Solow and Kirkwood (2002) attempted to motivate identity in three ways: through no interaction prior to the experiment (stranger treatment); through completing a group questionnaire (question treatment); using teams of participants drawn from a common community (i.e. members of the university marching band; community treatment). Significant differences decision-making were only identified in the community treatment. In perhaps the most comprehensive experimental study of identity in economic environments, Eckel and Grossman (2005) find that artificially motivated team identity (using many of the approaches discussed above) had no effect on cooperation in a series of public goods games. These results suggest that only when team members worked together to meet an initial goal was an identity effect discernable: cooperation with group members increased in this treatment, but by less than 10%.⁸ Other research using real social groups have identified

⁷In Rabbie and Horowitz (1969) prizes were allocated prior to gathering information on first impressions (which demonstrated an in-group bias). In Wit and Wilke (1992) uncertainty over rewards/prizes was resolved after the game was played.

⁸It is interesting to note the results of Oxoby and Spraggon (forthcoming, 2006) which suggest that fracturing identity in the lab may be easier than motivating identity: By differentiating participants in terms of endowments and earnings behavior, the behavioral effect of social preferences was significantly reduced (i.e. to levels predicted by Nash wealth maximization).

in-group favoritism. For example, Bernard et al. (2006) find that members of tribal groups in Papua New Guinea punish the unfair treatment of own-tribe members by outsiders more than the unfair treatment of non-tribe members. Similarly, Goette et al. (2006) finds evidence of greater cooperation in prisoners' dilemma games when individuals are interacting with members of their own group (platoons in the Swiss Army) and that third parties punish more when a member of their group suffers from defection. Finally, Hoff and Pandey (2006) identifies in-group favoritism in economic experiments when there are long-standing differences between groups (i.e. caste differences).

3 Experimental Design and Hypotheses

In this section we describe our experimental design and present a simple model of identity and social preferences to motivate our hypotheses. In each session, two groups (comprised of five to seven participants) met in separate rooms with one of two experimenters. Participants were informed that their group had been assigned to either Team 1 or Team 2 and presented with a task (emphasizing factor 1). All participants were informed that another team was a separate room and would be completing a similar task. Each experimenter read their group the instructions, and once all participants had acknowledged understanding of the instructions, the experiment began.⁹

Our main experiment occurred in three stages.¹⁰ In the first stage, each group was given 30 minutes to work together on a series of twenty questions. At the end of the 30 minutes, one answer sheet was collected from each group and scored based on the following schedule: if the group answered 10 or more questions correctly, each

⁹Instructions are presented in appendix A.

¹⁰The first two stages of each experiment were along the lines of the experiment discussed in Brown and Ross (1982).

team member would receive a lump-sum endowment of 100 lab dollars (1 lab dollar = \$0.10); if the group answered 9 or fewer questions correctly, each team member would receive 50 lab dollars. Our hope was that working together would establish a group identity among participants (through factor 2) which would be manifest in later stages of the experiment.

The type of questions answered differed between the groups. Team 1 was assigned 20 questions from the Graduate Records Examination (GRE) upon which their lump-sum endowment would be based. Team 2 was assigned a series of 20 animal congregation matching questions (e.g. matching gaggle to geese, murder to crows).¹¹ After the 30-minute identity-building task, each group was given a copy of the other group's quiz and informed of their score. Thus, after each group had completed their questions, they were informed of the type of task faced by the other group.

In the second stage of the experiment, both groups were brought into the experimental economic laboratory and divided in such a way that they were unable to see members of the other team. Participants were informed (via the computer stations) of the scores and endowments of both groups. After receiving this information, each participant was asked to gauge (on a scale of 1 to 10) the reasoning ability, creativity, verbal ability, motivation, and overall intelligence of the other group relative to their own. Our intent with this stage was to explore the effect of out-group opinions on in-group members (cf. Brown and Ross, 1982; Branscombe et al., 1999, 2002). Each group's average opinion of the other group was calculated and reported to the other group (i.e. the average opinions of Team t about Team s were reported to Team s , $s, t \in \{1, 2\}$, $t \neq s$). This procedure was meant to increase the degree of inter-group identity threat through the knowledge of out-group opinions.

Note that the use of different question types in the first stage (i.e. GRE vs. animal

¹¹Copies of both series of questions are available from the authors upon request.

matching) was designed to promote divergent, potentially illegitimate, opinions of the other group (emphasizing factor 3, see Branscombe et al., 2002). At first glance, the animal matching questions look easy to those who just answered GRE questions. The individuals may therefore consider answering 12 out of 20 animal matching questions as unimpressive and therefore rank the other team lower relative to themselves. On the other hand, individuals having answered the animal matching questions have a better idea of how challenging animal matching questions are and may therefore view the opinions of the other team as illegitimate.

The final stage of the experiment followed the proposer-responder games used by Andreoni et al. (2003) and Gneezy (2004): In this game, proposers chose a share of 24 lab dollars to keep for themselves and a share to allocate to a responder. Upon learning the amount received from the proposer, the responder was allowed to engage in costly punishment, decreasing the proposer's income by 1.5 lab dollars for each dollar paid. Both money from her endowment and money allocated by the proposer were available to engage in this costly punishment. Thus, the proposer's (P) and responder's (R) payoffs from a single round of decision making is given by

$$\Pi_P = (24 - x) - 1.5 \times y, \quad (1)$$

$$\Pi_R = x - y, \quad (2)$$

where $x \in [0, 24]$ (integer values) is the offer extended by the proposer and $y \geq 0$ (integer values) is the amount the responder chooses to punish. Individuals participated in six rounds of decision making with random (perfect stranger) matching in each round. To keep issues of identity salient in participants' decision making, participants were informed of their group and that of the other person with whom they were matched. Final payoffs to each participant were the sums of their final incomes in each round and their endowment from the first stage. Participants were

paid privately and in such a way so as they did not interact with members of the other group after receiving their payments.

We consider three treatments. The *opinion-punishment* (OP) treatment was conducted as described above. In the *no opinion treatment* (NO), the above protocols were followed with the exception of the portion of stage two in which opinions of the other group were elicited. This treatment serves as a control for the effect of out-group opinions on in-group and out-group interactions. In the *no punishment* (NP) treatment, responders did not have the opportunity to engage in negative reciprocity (i.e. reduce the income of proposers). As such, proposers engaged in six rounds of a dictator game with 24 lab dollars per round. This serves as a control for determining the extent to which the potential for negative reciprocity motivated offers in the OP treatment.

To motivate our hypotheses regarding the effects of identity on decision making in our experiment, consider the following simple model.¹² An agent i has consumption c_i with preferences represented by

$$U(c_i) = u(c_i) - \sigma(I_{ij})g(c_i - c_j) + \rho(I_{ij})\left(F_{ij} \times h^P(r_i) + (1 - F_{ij}) \times h^N(r_i)\right) \quad (3)$$

where c_j is the consumption of a referent agent j , $\sigma(I_{ij}) \geq 0$ is the weight assigned to differences in consumption and $\rho(I_{ij}) \geq 0$ it is the weight assigned reciprocal actions. We assume that $u(\cdot)$ is increasing and concave. The function increasing and convex function $g : \mathbb{R} \rightarrow \mathbb{R}$ measures how differences in consumption (inequities $c_i - c_j$) factor into utility. We assume that $g''(\cdot) > 0$ implying that the marginal disutility of having an additional dollar more when an endowment is split 55-45 is smaller than when

¹²This model contains aspects from other models of social preferences based on relative payoffs or inequity aversion (Bolton and Ockenfels, 2000; Charness and Rabin, 2002; Fehr and Schmidt, 1999), equity theory (Konow, 2000; Walster et al., 1978), reciprocity (Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006; Rabin, 1993), and the contextual aspects of interactions (e.g. Oxoby and Bischak, 2005). See Konow (2003) for a review of these various theories.

the endowment is split 70-30. Centrally, this captures the idea that if an individual dislikes favorable inequity, then the marginal cost of inequity should be larger when there is a greater disparity in payoffs.

We let F_{ij} be an indicator function taking on the value of 1 if agent j behaved fairly towards agent i and zero if agent j behaved unfairly towards agent i . The functions $h^P : \mathbb{R} \rightarrow \mathbb{R}$ and $h^N : \mathbb{R} \rightarrow \mathbb{R}$ measure how positive and negative reciprocity (r_i) affect utility. We assume that $h^P(0) = h^N(0) = 0$ with $h^P(r)$ and $h^N(r)$ increasing and concave in r . This implies that an individual derives positive utility from engaging in positive reciprocity when she has been treated fairly ($F_{ij} = 1$) and derives positive utility from negative reciprocity when she has been treated unfairly ($F_{ij} = 0$).¹³

To capture the role of identity in decision-making, we let I_{ij} denote an indicator function taking on the value of 1 if agents i and j identify with one another and zero otherwise. Given how behavioral prescriptions are affected by identity, we assume that $\sigma(1) > \sigma(0)$ and $\rho(1) > \rho(0)$. Thus, inequities in outcomes and reciprocal actions matter more when individuals identify with one another (i.e. are members of an in-group) than when they do not identify with each other (i.e. are not members of the same in-group).

We now consider the analysis of our experiment within the context of this model. First, we consider how offers in the absence of punishment are affected when the proposer ($i = 1$) identifies with the receiver ($j = 2$). Note that a proposer's utility maximizing offer satisfies

$$-u' + 2\sigma(I_{12})g' \leq 0, \tag{4}$$

since $c_1 = \omega - x$ and $c_2 = x$ where ω is the proposer's endowment and x is the offer

¹³A similar approach is captured in the kindness functions used in the reciprocity models of Rabin (1993) and Dufwenberg and Kirchsteiger (2004).

extended. At an interior solution, it is straightforward to show that

$$\frac{\partial x_1}{\partial \sigma} = \frac{-2g'}{u'' - 4\sigma(I_{12})g''} > 0. \quad (5)$$

Thus if proposer identifies with the receiver (and hence $\sigma(1) > \sigma(0)$) she will extend larger offers than when she interacts with someone with whom she does not identify. In terms of our experiment, we have the following hypothesis:

Hypothesis 1 *Group Identity and Inequity:* *Individuals will cooperate more with individuals with whom they identify than with individuals with whom they do not identify. That is, proposers will extend larger offers to members of the in-group than to members of the out-group.*

Experimentally, this hypothesis follows from past research on group identity and group processes (e.g. Nass et al., 1996; Turner, 1982).

Further, more recent research (Branscombe et al., 1999, 2002) argues that negative (and potentially illegitimate) out-group opinion reinforces group identification. In terms of the model (equation 5) these out-group opinions should increase the degree to which individuals identify with members of their in-group and we have the following:

Hypothesis 2 *Cooperation and the Degree of Identification:* *Lower out-group opinion (relative to own opinion) will reinforce in-group identity, resulting in greater cooperation within the in-group. Higher out-group opinion (relative to own opinion) will reduce the degree of in-group identification, resulting in less cooperation within the in-group (relative to the out-group).*

Thus, negative out-group opinion (relative to in-group opinion) will reinforce group identification, thereby increasing offers to members of the in-group relative to the out-group. Similarly, positive out-group opinion (relative to in-group opinion) will result in group disassociation and, as such, less in-group cooperation. Further,

the more divergent are these opinions, the larger will be their effects on in-group cooperation (Brown and Ross, 1982).

Since receivers in our experiments are restricted to costly punishment, $r_2 < 0$. Thus we focus on negative reciprocity and how a receiver may punish a proposer in the face of an unfair offer (i.e. $F_{21} = 0$). Given an unfair offer from agent 1, agent 2 will choose the level of punishment r_2 such that

$$-u' + \rho(I_{21})\pi(h^N)' \leq 0, \quad (6)$$

where π is the productivity of punishment (in our design $\pi = 1.5$). It is straightforward to show that

$$\frac{\partial r_2}{\partial \rho} = \frac{\pi(h^N)'}{u'' + \rho(I_{21})\pi^2(h^N)''} > 0. \quad (7)$$

Thus, in response to an unfair offer from agent 1, agent 2 will engage in greater negative reciprocity (here, punishment) if she identifies with agent 1. Such behavior is similar in spirit to the ideas discussed in Akerlof and Kranton (2000) and Stroebe et al. (2005) in terms of identity threat, the violation of behavioral prescriptions, and identity-based reciprocity.

From this we have the following hypothesis:

Hypothesis 3 *Group Identity and Punishment:* *In response to unfair offers, individuals will engage in greater punishment when interacting with members of the in-group than with members of the out-group.*

Essentially, receiving an unfair offer from an individual with whom one identifies violates the implicit behavioral norms associated with a shared identity (in our environment). As such an individual is more likely to punish, and to punish more, unfair offers coming from individuals with whom identity is shared (Hypothesis 3). This effect will be stronger in contexts in which group identity is stronger, for example, as

is the case when there exist negative out-group opinions. In contrast, when an individual does not identify with another, there may be no implicitly assumed behavioral norm of fairness. As such unfair offers may be more expected (and hence tolerated) when they come from someone with whom one does not identify. This hypothesis is related to the ideas discussed by Akerlof and Kranton (2000, 2005) regarding threat to one's identity: We posit that lack of cooperation (i.e. an unfair offer) by an in-group member imposes an externality in the form of intra-group identity threat. Given this violation of the behavioral prescription dictated by being a member of the in-group, an individual will be more likely to engage in negative reciprocity in response to unfair actions from in-group members relative to those from out-group members.

4 Results

Participants were recruited from the student body at our university. The experiments were conducted in the experimental economics laboratory and programmed in z-Tree (Fischbacher, 1999). Over the course of the study, we conducted eight sessions of the OP treatment (91 participants), four sessions of the NO treatment (37 participants), and four sessions of the NP treatment (38 participants).

We begin by considering the behavior of proposers. We let the variable *Invest* measure the amount invested by each proposer in each round of the game. In order to investigate how cooperation differed between in-group and out-group pairs in each treatment, we use a fixed-effects regression of *Invest* on responder's status as a member of the in-group (*SameGroup* = 1) or the out-group (*SameGroup* = 0). Results for each treatment are presented in Table 1.

The coefficients on *SameGroup* are all positive and statistically significant, evidence that in all of the treatments interactions with in-group members elicited larger offers than did interactions with out-group members. As such, we are unable to reject

Table 1: Proposers' Behavior

OP Treatment	Coefficient	(Std. Err.)
SameGroup	4.067**	(0.618)
N		271
R ²		0.627
F _(89,181)		3.415
NO Treatment	Coefficient	(Std. Err.)
SameGroup	2.324**	(0.867)
N		113
R ²		0.661
F _(34,78)		4.464
NP Treatment	Coefficient	(Std. Err.)
SameGroup	2.490 [†]	(1.344)
N		114
R ²		0.437
F _(38,75)		1.533
Significance levels : † : 10% * : 5% ** : 1%		

Hypothesis 1. Thus, proposers appear to have cooperated more with responders of their in-group. Further, using Chow tests we are unable to reject the hypotheses that $SameGroup_{NO}=SameGroup_{OP}$ ($p = 0.165$) and $SameGroup_{NP}=SameGroup_{OP}$ ($p = 0.203$). That is, offers do not significantly differ between treatments where there is a punishment stage and treatment where there is not a punishment stage. Similarly, offers do not differ significantly between treatments where out-group opinion is present and treatments where out-group opinion is absent. This means that out-group opinions did not provide additional motivation for identity (manifesting itself in larger offers). Strikingly, punishment did not motivate greater offers (OP treatment relative to NP treatment) as evidenced in other experiments. For example, Fehr and Gächter (2000, 2001) find that the potential for punishment increases contributions to public goods and gift-exchange.

To explore Hypothesis 2, we need concentrate on the difference between an individual's opinion of the relative standing of the two groups and the out-group's average opinion of the relative standing. To do this, we created two new variables based on the differences between these opinions: *NegDiff* represents the degree to which the other group ranked the relative standing of the individual's own group more highly than the individual did (i.e. the out-group had a relatively more positive opinion of the in-group), while *PosDiff* represents the degree to which the other group ranked the relative standing of the individual's own group less highly than the individual did (i.e. the out-group had a relatively more negative opinion of the in-group). We use these variables to construct two interaction variables, $SameGroup * NegDiff$ and $SameGroup * PosDiff$.

The results of this regression are reported in table 2. Note that the results from both the OP and the NP treatments are comparable in both magnitude and sign. In both treatments, the coefficient on $SameGroup * PosDiff$ is positive and significant, implying that if the out-group had a lower opinion of the in-group than the individual

did, norms of cooperation were reinforced in interactions with in-group members. Further, in the OP treatment the coefficient on *SameGroup * NegDiff* is negative and significant.¹⁴ This is evidence that if the out-group had a higher opinion of the in-group than the individual did, norms of cooperation were weakened within the in-group and offers to in-group members were correspondingly lower.

Summarizing, lower out-group opinion (relative to own opinion) led to a strengthening of group identity and thus, adherence to group norms, while higher out-group opinion (relative to own opinion) led to group disassociation and a weakening of group norms. This conforms with our Hypothesis 2: Relatively lower opinion conferred on a group via illegitimate means leads to a strengthening of group cooperation norms, while relatively higher opinion conferred on a group led to a weakening of group cooperation norms. This result matches those of Branscombe et al. (1999) in which illegitimate negative out-group opinions contributed to in-group identification.

We next turn to the manner in which offers were reciprocated by responders. The variable *Punish* measures the amount paid by responders to decrease the incomes of proposers. Note that Nash wealth maximization implies responders choose *Punish* = 0. The proportions of respondents that chose *Punish* = 0 in the OP and NO treatments are presented in Table 3, both for all offers and for offers less than 50% (i.e. less than 12 lab dollars). Note that a significant proportion of participants in the OP and NO treatments chose non-zero punishment amounts. Furthermore, in the OP treatment responders punished unfair offers made by members of their in-group more often than those from members of their out-group. Second, the table clearly shows the effect of out-group opinion on punishment behavior in in-group pairs: The treatment with out-group opinion had a larger proportion of participants choosing non-zero punishment amounts.

In the OP treatment, the distributions of punishment within in-group pairs and

¹⁴The coefficient on *SameGroup * NegDiff* is negative, but insignificant in the NP treatment.

Table 2: Investment Behavior by Proposers, Controlling for Opinion

OP Treatment	Coefficient	(Std. Err.)
SameGroup*NegDiff	-0.091**	(0.019)
SameGroup*PosDiff	0.346**	(0.100)
N		271
R ²		0.613
F _(90,180)		3.167
NP Treatment	Coefficient	(Std. Err.)
SameGroup*NegDiff	-0.276	(0.415)
SameGroup*PosDiff	0.459 [†]	(0.271)
N		114
R ²		0.437
F _(39,74)		1.47
Significance levels : † : 10% * : 5% ** : 1%		

Table 3: Proportion of Nash Behavior by Responders to All Offers and to Unfair Offers ($Invest < 12$), By Treatment

Pairing	OP	NO
In-Group	68	65
Out-Group	70	59
In-Group (unfair offer)	34	50
Out-Group (unfair offer)	67	56

within out-group pairs in response to unfair offers are significantly different (Wilcoxon $p < 0.01$). The kernel densities of these two distributions are presented in Figure 1. Note the distinct difference between responders' behavior in in-group pairs and out-group pairs: Given an unfair offer, a large mass of responders in in-group pairs chose a positive punishment amount, while most responders in out-group pairs chose very small (or zero) punishment amounts.

It is not only with respect to the proportion of Nash behavior that punishment within in-group pairs in the two treatments differ: The distributions of punishment within in-group pairs in response to unfair offers are significantly different between the treatments with and without out-group opinion (Wilcoxon $p = 0.10$). The kernel densities of these two distributions are presented in Figure 2. Note the larger mass of non-zero punishment choices when participants interacted with in-group members in the OP treatment. Interestingly, the distributions of punishment within out-group pairs are not significantly different between the OP and NO treatments (Wilcoxon $p = 0.41$), suggesting that out-group opinion does nothing to affect interactions between groups.

A further examination of the determinants of the punishment behavior can be conducted through a fixed effects regression using the data from the OP treatment. Note that we limit observations in the dataset to those where offers were strictly less than 12 lab dollars (i.e. unfair offers). The results of this regression are reported in Table 4. First, the significant positive coefficient on *SameGroup* shows that, on average, responders punished proposers from their in-group more than proposers from their out-group. Further, as evidenced by the significant negative coefficient on *Invest * SameGroup* and the insignificant coefficient on *Invest*, responders' punishment behavior was much more sensitive to the investment amounts from in-group members than out-group members.

These results conform with our Hypothesis 3: Individuals engaged in more pun-

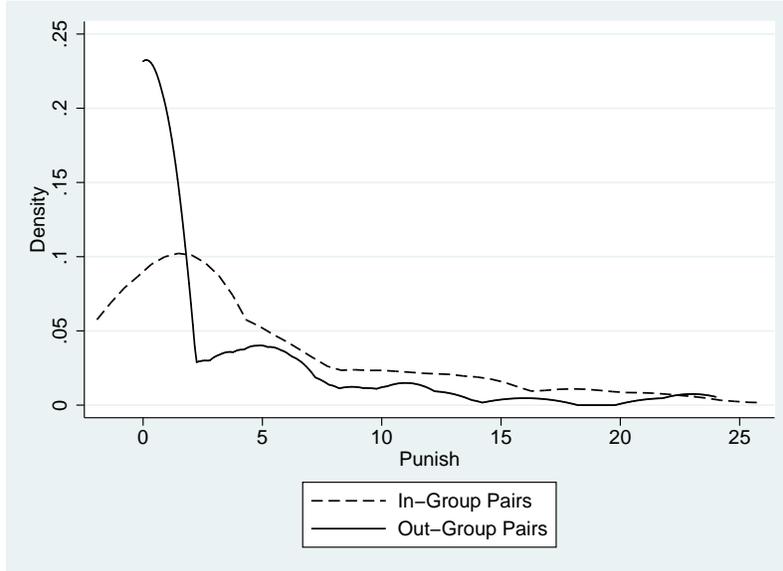


Figure 1: Distributions of Punishment to Unfair Offers, by Pairing

ishment in response to uncooperative behavior by those in their in-group. Further, given the differences in Nash behavior shown in Table 3, the presence of out-group opinion strengthened this effect: In-group punishment towards unfair offers seems to be higher in the OP treatment than the NO treatment. That is, responses to intra-group identity threat appear to be higher in the presence of inter-group identity threat.

Table 4: Punishment Behavior by Responders to Unfair Offers (i.e. $Invest < 12$) in OP Treatment

Variable	Coefficient	(Std. Err.)
Invest	-0.211	(0.224)
SameGroup	5.231**	(1.887)
Invest*SameGroup	-0.580 [†]	(0.304)
N		154
R ²		0.542
F (79,74)		1.11
Significance levels : † : 10% * : 5% ** : 1%		

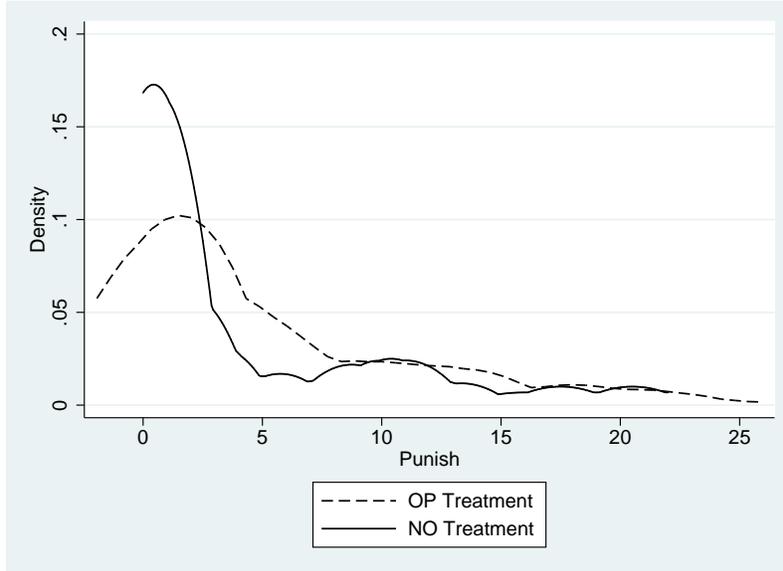


Figure 2: Distributions of Punishment to Unfair Offers Among In-Group Pairs, By Treatment

5 Conclusion

In many of the interactions in which individuals engage, elements of identity influence decision making: The ways individuals vote, offer assistance, and locate themselves are all influenced by the degrees of identification they experience with others. In this paper, we investigate the determinants of group identity and its effects in a series of simple two-person bargaining games. These experiments permit us to explore two aspects of identity and identity threat: inter-group identity threat (motivated through out-group opinions) and intra-group identity threat (made salient through the violation of tacit in-group behavioral norms).

We find that identity threat, in both of its forms, has a significant effect on behavior. Our results suggest that *(i)* individuals cooperate more with members of their in-group, and *(ii)* in-group cooperation is strengthened by relatively lower out-group opinion and weakened by relatively higher out-group opinion. We also find that individuals engage in costly punishment more often with in-group members in

response to behaviors which violate tacit in-group norms of cooperation.

We hypothesize that behavior stemming from either type of threat is caused by the need to defend a positive self-concept, which is itself tied to the identity of the group (Tajfel and Turner, 1986), pointing to an interesting aspect of punishment among individuals who share identity: Unlike previous studies in which negative reciprocity motivates greater cooperation (e.g. Fehr and Gächter, 2000, 2001; McCabe et al., 2003), punishment has no such effect in our experiments. As such, our results suggest that negative reciprocity may be a means of reaffirming one's in-group identity in the face of a within group offer which is deemed unfair. This character of response is similar to the cases of "acting out" described in Akerlof and Kranton (2000, 2005) in which individuals engage in costly actions in order to reenforce their own identities.

References

- Akerlof, G. A., Kranton, R. E., 2000. Economics and identity. *Quarterly Journal of Economics* 115 (3), 715–758.
- Akerlof, G. A., Kranton, R. E., 2002. Identity and schooling: Some lessons for the economics of education. *Journal of Economic Literature* 40 (4), 1167–1201.
- Akerlof, G. A., Kranton, R. E., Winter 2005. Identity and the economics of organizations. *Journal of Economic Perspectives* 19 (1), 9–32.
- Alesina, A., Baqir, R., Easterly, W., 1999. Public goods and ethnic divisions. *Quarterly Journal of Economics* 114, 1243–1284.
- Alesina, A., Ferrara, E., August 2000. Participation in heterogeneous communities. *Quarterly Journal of Economics* 115 (3), 847–904.
- Alesina, A., Glaeser, E. L., Sacerdote, B., Fall 2001. Why doesn't the U.S. have a European style welfare state? *Brookings Papers on Economic Activity*, 187–278.
- Andreoni, J., Harbaugh, W., Vesterlund, L., 2003. The carrot or the stick: Rewards, punishment and cooperation. *American Economic Review* 93 (3), 893–902.
- Badgett, M. V. L., 1995. The wage effects of sexual orientation discrimination. *Industrial Labor Relations Review* 48 (4), 726–739.

- Berg, N., Lein, D., 2002. Measuring the effect of sexual orientation on income: Evidence of discrimination? *Contemporary Economic Policy* 20 (4), 394–414.
- Bernard, H., Fehr, E., Fischbacher, U., 2006. Tribal identity and altruistic norm enforcement. University of Zurich working paper.
- Bolton, G. E., Ockenfels, A., 2000. A theory of equity, reciprocity, and competition. *American Economic Review* 30 (1), 166–193.
- Branscombe, N. R., Schmitt, M. T., Harvey, R. D., 1999. Perceiving pervasive discrimination among African Americans: Implications for group identification and well-being. *Journal of Personality and Social Psychology* 77 (1), 135–149.
- Branscombe, N. R., Spears, R., Ellemers, N., Doosje, B., September 2002. Intragroup and intergroup evaluation effects on group behavior. *Personality And Social Psychology Bulletin* 28 (9), 744–753.
- Brown, R. J., Ross, G. F., 1982. The battle for acceptance: An investigation into the dynamics of group behavior. In: Tajfel, H. (Ed.), *Social Identity and Intergroup Relations*. Cambridge University Press, Cambridge.
- Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. *Quarterly Journal of Economics* 117 (3), 817–869.
- Coate, S., Loury, G. C., 1993. Will affirmative-action policies eliminate negative stereotypes? *American Economic Review* 83 (5), 1220–1240.
- Devine, P. G., Plant, E. A., Harrison, K., April 1999. The problem of “us” versus “them” and AIDS stigma. *American Behavioral Scientist* 42 (7), 1212–1228.
- Dufwenberg, M., Kirchsteiger, G., 2004. A theory of sequential reciprocity. *Games and Economic Behavior* 47 (2), 268–298.
- Durlauf, S., 1999. The case against social capital. *Focus* 20 (3), 1–4.
- Eckel, C. C., Grossman, P. J., 2005. Managing diversity by creating team identity. *Journal of Economic Behavior and Organization* 58, 371–392.
- Falk, A., Fischbacher, U., 2006. A theory of reciprocity. *Games and Economic Behavior* 54 (2), 293–315.
- Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods experiments. *American Economic Review* 90 (4), 980–994.
- Fehr, E., Gächter, S., 2001. Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives* 14 (3), 159–181.
- Fehr, E., Schmidt, K., 1999. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114 (3), 817–868.

- Fischbacher, U., 1999. Toolbox for readymade economic experiments. Tech. Rep. IEW Working paper 21, University of Zurich.
- Gerber, G., 1998. Status in same-gender and mixed-gender police dyads: Effects on personality attributions. *Social Psychology Quarterly* 59, 350–363.
- Glaeser, E. L., Laibson, D., Sacerdote, B., 2002. An economic approach to social capital. *The Economic Journal* 112, F437–F458.
- Gneezy, U., 2004. The W effect of incentives. University of Chicago working paper.
- Goette, L., Huffman, D., Meier, S., 2006. The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups. Institute for Empirical Research in Economics.
- Haslam, S. A., 2001. *Psychology in Organizations: The Social Identity Approach*. Sage Publications, London.
- Hoff, K., Pandey, P., 2006. Caste identity, belief systems and durable inequalities. World Bank Policy Research working paper 3351.
- Konow, J., 2000. Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review* 90 (4), 1072–1091.
- Konow, J., 2003. Which is the fairest one of all? A positive analysis of justice theories. *Journal of Economic Literature* 41 (4), 1186–1237.
- Kramer, R. M., Pommerenke, P., Newton, E., December 1993. The social context of negotiation: Effects of social identity and interpersonal accountability on negotiator decision making. *Journal of Conflict Resolution* 37 (4), 633–655.
- Lee, W., Roemer, J. E., 2004. Racism and redistribution in the United States: A solution to the problem of American exceptionalism. Cowles Foundation Discussion Paper No. 1462.
- Lundberg, S. J., Startz, R., 1983. Private discrimination and social intervention in competitive labor markets. *American Economic Review* 73 (3), 340–347.
- Mackie, D. M., 1986. Social identification effect in group polarization. *Journal of Personality and Social Psychology* 50 (4), 720–728.
- Mackie, D. M., Worth, L. T., Asuncion, A. G., 1990. Processing of persuasive in-group messages. *Journal of Personality and Social Psychology* 58 (5), 812–822.
- Matheson, K., Cole, B. M., 2004. Coping with a threatened group identity: Psychosocial and neuroendocrine responses. *Journal of Experimental Social Psychology* 40, 777–786.

- McCabe, K. A., Rigdon, M. L., Smith, V. L., 2003. Positive reciprocity and intentions in trust games. *Journal of Economic Behavior and Organization* 52 (2), 267–275.
- Nass, C., Fogg, B., Moon, Y., 1996. Can computers be teammates? Affiliation and social identity effects in human-computer interaction. *International Journal of Human-Computer Studies* 45 (6), 669–678.
- Oxoby, R. J., 2004. Cognitive dissonance, status, and growth of the underclass. *The Economic Journal* 114, 729–749.
- Oxoby, R. J., 2005. Promoting an inclusionary economy. Social Development Canada issues paper.
- Oxoby, R. J., Bischak, D., 2005. Passing the time: Other-regarding behavior and the sunk cost of time. University of Calgary, Department of Economics discussion paper 2005-03.
- Oxoby, R. J., Spraggon, J., 2006. A clear and present minority: Heterogeneity in the source of endowments and the provision of public goods. Institute for Advanced Policy Research working paper.
- Oxoby, R. J., Spraggon, J., forthcoming. Yours and mine: Property rights in dictator games. *Journal of Economic Behavior and Organization*.
- Putnam, R. D., 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press, Princeton, NJ.
- Rabbie, J. M., Horowitz, M., 1969. Arousal of ingroup-outgroup bias by a chance win or loss. *Journal of Personality and Social Psychology* 13 (3), 269–277.
- Rabin, M., 1993. Incorporating fairness into game theory and economics. *American Economic Review* 83 (5), 1281–1302.
- Scheepers, D., Ellemers, N., 2005. When the pressure is up: The assessment of social identity threat in low and high status groups. *Journal of Experimental Social Psychology* 41, 192–200.
- Solow, J. L., Kirkwood, N., 2002. Group identity and gender in public goods experiments. *Journal of Economic Behavior and Organization* 48, 403–412.
- Steele, C. M., Aronson, J., 1995. Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology* 69, 797–811.
- Stroebe, K., Lodewijkx, H. F., Spears, R., 2005. Do unto others as they do unto you: Reciprocity and social identification as determinants of ingroup favoritism. *Personality and Social Psychology Bulletin* 31 (6), 831–845.

- Tajfel, H., Billig, M., Bundy, R., Flament, C., 1971. Social categorization in intergroup behavior. *European Journal of Social Psychology* 1, 149–178.
- Tajfel, H., Turner, J., 1986. *The social identity theory of intergroup behavior*. Nelson-Hall, Chicago, pp. 7–24.
- Turner, J. C., 1982. Towards a cognitive redefinition of the social group. In: Tajfel, H. (Ed.), *Social Identity and Intergroup Relations*. Cambridge University Press, Cambridge.
- Walster, E., Walster, G., Berscheid, E., 1978. *Equity: Theory and Research*. Allyn and Bacon, Boston.
- Wann, D. L., Grieve, F. G., 2005. Biased evaluations of in-group and out-group spectator behavior at sporting events: The importance of team identification and threats to social identity. *Journal of Social Psychology* 145 (5), 531–545.
- Wilder, D. A., 1990. Some determinants of the persuasive power of in-group and out-groups: Organization of information and attribution of independence. *Journal of Personality and Social Psychology* 59, 1202–1213.
- Wilder, D. A., Shapiro, P. N., 1984. Role of out-group cues in determining social identity. *Journal of Personality and Social Psychology* 47 (2), 342–348.
- Wit, A. P., Wilke, H. A. M., 1992. The effect of social categorization on cooperation in three types of social dilemmas. *Journal of Economic Psychology* 13, 135–151.
- Zak, P. J., Knack, S., 2001. Trust and growth. *The Economic Journal* 111 (470), 295–321.

A Instructions

These instructions are those used in the OP and NO treatments. The opinion questions were not mentioned in the instructions in an effort to avoid the responses being used strategically or more calculated responses regarding one's opinion of the other group.

This is an experiment in decision making. During this experiment you will make a number of decisions. These decisions will result in a payoff which will be paid in cash. Your payment for this session is both compensation for your time as well as for the effort you put into making your decisions.

This experiment will consist of two parts. During the first part of the experiment, you will work with the group of individuals around you. The details of this part of the experiment will be described below.

The second part of the experiment will be conducted in the experimental economics laboratory and will consist of 6 rounds of decision making. During this part of the experiment we ask that you refrain from speaking with others. If you have any questions, please raise your hand and an experimenter will assist you.

Each of the 6 rounds will proceed as follows:

1. In each round you will be randomly matched with a different person. You will be notified of the team this person belongs to. One of you will be randomly assigned to the role of Proposer and the other will be randomly assigned to the role of Responder.
2. Each round takes place in two stages. In the first stage the Proposer is asked to propose a split of 24 lab dollars between him/herself and the Responder. Note that all amounts in the experiment are denominated in lab dollars. At the end of the experiment these will be converted into Canadian dollars at a rate of 1 lab dollar equal to \$0.10 Canadian dollars.
3. The Proposer's choice is then revealed to the Responder. The Responder determines how to respond to the proposed allocation. They may decrease or make no change to the Proposer's income. In order to adjust the Proposer's income, the Responder must pay a cost: For every lab dollar the the Responder pays, the Proposer's income will decrease by 1.5 lab dollars. The money to pay for the decrease will come from the money allocated to the Responder. If the amount paid to decrease the Proposer's income is in excess of what was allocated to the Responder, this money will be taken out of the Responder's endowment. Hence, the cost is not limited to the amount the Proposer allocated to the Responder.
4. After everyone has made their decisions, you will be notified of your final income from the round and that of the person with whom you were matched. After viewing this information, please click the OK button. You will then be randomly matched with another individual for the next round.
5. Your total payoff from the experiment will be the sum of your final incomes in each round.

Before you begin these decision making rounds, you will work as a team to determine your endowment. You have been assigned to Team 2, consisting of the people in the room with you. Your team's task in the first part of the experiment is to answer a series of 20 questions. You will have 30 minutes to work together on this task. At the end of the 30 minutes, the answer sheet will be collected from you and scored. Only one answer sheet should be completed within your team.

The endowment all members of your team will receive will depend on the number of questions your team correctly answers. Specifically, if your team answers 10 or more questions correctly, each team member will receive 100 lab dollars as a lump-sum endowment. If your team answers 9 or fewer questions correctly, each team member will receive 50 lab dollars as a lump-sum endowment.

Note that there is another team of people, Team 1, in a separate room. They too have received these instructions and will have 30 minutes to complete 20 questions.

After the exam has been scored, you will be brought into the same room and divided in such a way that you will not be able to see the members of the other team. Both teams will participate in the 6 decision making rounds described above.

At the end of the experiment, you will be asked several questions. All of this information is confidential. Upon completion, you will be paid in private and each team will leave at separate times so as to preserve anonymity.

Once everyone has had an opportunity to ask any final questions we will begin the experiment.

Warm-up Exercises

1. Suppose your team answers 12 questions correctly. How much will each member's endowment be?
2. Suppose that you have an income of 4 lab dollars in a decision making round and the other person has an income 20. If you pay 2 to reduce the other person's income, what will be your final income in the round? What will be the final income of the other person?