Individual Behavior and Group Membership[1]

First Draft: December 24, 2004

This Draft: March 4, 2005

Gary Charness
Department of Economics
UCSB

Luca Rigotti
The Fuqua School of Business
Duke University

Aldo Rustichini
Department of Economics
University of Minnesota

Abstract

People who are members of a group, and identify with it, behave differently from people in isolation. The way in which the behavior differs depends in subtle ways from the way in which the nature of the group is perceived, as well as its saliency, and also on the way in which people perceive that the behavior of others is affected by the group.

We study these hypotheses in a strategic experimental environment. Participants are allocated randomly to two groups (Row and Column players), and a room is assigned to each group. The saliency of the group membership is manipulated by making the group present as an audience in the corresponding room, or not. We use two stage games, the Battle of the Sexes and Prisoner's Dilemma.

We show that the saliency of the group affects behavior of members, as well as the behavior of people in the other group, and that participants anticipate these effects. Group membership increases the aggressive stance of the hosts (people who have their group members in the audience). The effect on the outcomes of this increased aggressive stance depends on the game: In the Battle of the Sexes, the aggressiveness of hosts leads to coordination on an efficient, alternating outcome; in the Prisoner's Dilemma, it leads to conflict and inefficient outcomes.

[1] Gary Charness, Dept. of Economics, UCSB, 2127 North Hall, Santa Barbara, CA 93106-9210, 805-893-2412, charness@econ.ucsb.edu; Luca Rigotti, Fuqua School of Business, Duke University, Box 90120, Durham, NC 27708, 919-660-7780, rigotti@duke.edu; Aldo Rustichini, Dept. of Economics, University of Minnesota, 1035 Heller Hall, 271 19th Avenue South, Minneapolis, MN 55455, 612-625-4816, arust@econ.umn.edu

The research was supported in part by research funds from UCSB, research funds from Duke University, and the NSF grant NSF/SES-0136556 to Aldo Rustichini.

1. INTRODUCTION

Do players who are members of a group, and identify with it, behave differently in strategic environments than players in isolation? While game theory and economics are very effective tools for analyzing behavior in social interaction, they are typically silent on the effect of group membership on behavior. Nevertheless, we feel that there may very well be effects from group membership on the decisions made by individual members. If group membership affects behavior, then the conclusions of a theory that is based on the assumption that behavior (even altruistic, reciprocal, or fair behavior) is always simply determined at the individual level should be revised.

The aim of our study is to begin a systematic analysis of group behavior. The basic idea is that, when groups are involved, human behavior is better understood as social behavior. However, we depart from social psychology in that we consider strategic environments and we wish to trace how and why this difference arises. If we can understand how group members (and outsiders) modify their behavior compared to individuals acting in isolation, we can provide a theory of how groups influence strategic behavior in economic contexts.

We test this general hypothesis in an experimental setup by examining whether making the presence of the group salient affects behavior in two very different games, the Battle of the Sexes and the Prisoner's Dilemma; we also test the additional hypothesis that the saliency of the group not only affects the behavior of group members, but also the perception that players have of the effect of group membership on the behavior of others. Our games are played face-to-face, but without any verbal communication; we vary whether an audience comprised of members of one group is in attendance and whether feedback about the decisions made is immediately given. We impose some group commonality by having the audience share in the payoffs.

Our main result is that there is a substantial and significant effect of group membership on behavior, particularly when an audience is present and feedback is provided. We find that players are much more aggressive when they play in front of their informed partisans. This change in behavior has very different effects in our games: While this mechanism appears to serve as an equilibrium-selection device in the Battle of the Sexes, improving coordination and social payoffs, it has an adverse effect on cooperation in the Prisoner's Dilemma, with social payoffs correspondingly diminished. This result has implications for a broad range of social and behavioral sciences, so that it is important to understand their behavioral underpinnings. Our data offer insights into how group behavior operates.

An audience is introduced in our experiment to manipulate the saliency of the group, but the presence of an audience is a common occurrence, and the results are interesting from this point of view as well. In most jobs, we conduct tasks while co-workers can observe us. While the extent of this observation depends on how open the working environment is, a natural conjecture is that peer pressure to perform is effective, and is higher in a more open environment. In a similar vein, children in a classroom perform their daily activity in front of the audience of their classmates, while home-schooled children do not. Law firms are known to bring a larger number of lawyers to a meeting than needed for bargaining; this sympathetic audience may intimidate the opposing side. Professional athletes have spectators watching their performances; the home-field advantage shows how important a supportive audience can be.

The issue of the determinants of group behavior is huge. Some of the implications are well-known and classic topics in a long-running debate between economics and sociology. A clear example is the alternative between the scientific-management method of the organization of labor (Taylor 1911) and the human-relations paradigm (Mayo 1949). Some commentary is more

recent: Huntington proposes, on p. 20 of *The Clash of Civilizations*, the thesis that "culture and cultural identities, which at the broadest level are civilization's identities, are shaping the pattern of cohesion, disintegration and conflict in the post-Cold War world."[2]

Traditional game theory offers no explanation for the observed effects, if we assume that the payoffs indeed fully represent the players' utilities and that there are no asymmetric expectations. Nevertheless, if we permit the players' expectations to be influenced by the presence of a partisan audience or consider that the utility of the audience may be internalized asymmetrically, game theory could address our results. We discuss how these expectations might form, building on previous work in social psychology and economics regarding group membership, identity, and performance.

The remainder of the paper is organized as follows: We review the relevant literature in Section 2, discussing how the theories and models therein might apply to our context. We describe our experiments in detail in Section 3, and present the experimental results in Section 4. We discuss our results and conclude in Section 5.

2. SOCIAL IDENTITY AND CATEGORIZATION

In the economics literature, Akerlof and Kranton (2000) have advanced the view that identity, or a person's sense of self, can affect economic outcomes such as gender discrimination in the workplace, poverty and social exclusion, and the household division of labor. Our results may stem from a specific aspect of group identity, wherein the physical presence of the group can itself have the effect of fostering this identity.

[2] On p. 21, he continues: "In the post-Cold War world, the most important distinctions among people are not ideological, political, or economic. They are cultural. People and nations are attempting to answer the most basic question humans can face: Who are we?''

There is a large literature in social psychology on in-group and out-group behavior, or *social categorization*. The initial and fundamental question that this research addressed is the origin of discrimination: Why do people behave in a way that is favorable to members of their own group, at a cost to the out-group? The seminal paper in this area is Tajfel, Billing, Bundy & Flament (1971). This study finds that adolescent boys favor members of their own experimental in-group in an allocation task, even in a situation devoid of the usual trappings of in-group membership.

The experimental method is driven by the search for the *minimal group*, or the weakest cohesion that will produce discriminatory behavior. The issue is significant because if discrimination arises in groups with only weak ties, then this is evidence that group behavior is a generic human instinct, rather than the exclusive outcome of deeply-rooted traditions, religion or blood. Tajfel *et al.* (1971) specified a set of criteria required for a group classification to be minimal. The conditions are: (1) No face to face interaction, (2) complete anonymity of group membership, (3) no instrumental link between group categorization and the nature of the responses required to the subjects, (4) no difference across choices in the material payoffs for the chooser, (5) competition between group motivation and some other motivation, and (6) the decision should be important to the subjects.[3]

The basic experiment used in this study (and typically in the others that followed) consists of a group allocation rule and a task. The group allocation rule creates two distinct groups out of an otherwise homogeneous set of participants. In one treatment, people guessed the number of dots present in an image flashed on a screen, while in another treatment people expressed preferences between pairs of abstract paintings by Klee and Kandisky. In each case,

[3] Note that our experiment violates the first five requirements, so it is not a minimal-group experiment.

participants were told that people with similar guesses or preferences were linked to form two separate groups, although in reality the matching was random. The task consisted of choosing an allocation of money for two other people, who are each identified as being in-group or out-group members, from a set of many feasible pairs of allocations. These were chosen so that both the total amount and the share to the two subjects varied systematically.

There are two fundamental findings in this literature. The first is that the cohesion of minimal group may be very weak, and discrimination still occurs: In Billig and Tajfel (1973) subjects are allocated to groups randomly, and subjects know this; even so, the effect (discrimination) is strong and significant. The second result is that discrimination has a very distinct objective: Making the outcome for the own group better than the outcome of the opponent group. The vector of feasible options was selected so that comparisons among different explanations for the behavior were possible, such as the maximum joint payoff, fairness, and the maximum difference between in-group and out-group allocations. The feature that was predominantly (and in most studies exclusively) significant was the maximum difference, even when this lowered the payoff of an in-group member.

The theory of social categorization offers an explanation for this observed behavior. The theory provides a foundation for social identity by trying to explain how subjects move from the perception of themselves as individuals to the perception of themselves as part of a group. First, people are presumed to strive for a sense of social identity and positive self-image.[4] In Tajfel and Turner (1979) the authors point out that the image of the group or groups to which they belong is an important component of this self image; thus people favor, and whenever possible try to produce, a favorable comparison between their in-group and out-group. Hence the attempt

[4] Huntington (1966) suggests that this need for individuals to hold an identity as one of the original motivations for civilizations, and the need for distinctiveness as a fundamental reason for their clash.

of subjects in the minimal group experiment to discriminate in favor of the in-group, by increasing the difference in outcomes in favor of their group, as the choice that among all produces the largest comparative gain to the own group.

In a world of imperfect information, this may be reasonable behavior. Suppose that there is an unobservable variable (skill) that is different across individuals, but may be correlated across members of a group. An outside observer can use the information on the performance of one of the members of a group to make inferences about the distribution of the skill in the group. If the realization of the variable is good, and the correlation positive, a member of the group may want an outside observer to have a positive opinion of the group because this will induce also a positive opinion of the member.

In our view, implicit in this theory are three channels by which membership in a group can affect behavior. First, in accordance with the comparative argument presented above, we presume that people use membership in a group as a source of information for themselves and a signal to others of their social identity. This leads to discrimination in favor of one's in-group, with the effect strongest when group cohesion or identification is high. Second, whenever there is a conflict among individuals on the allocation of resources, membership into a group can help the individual in the conflict. In this sense, membership in a group confers a sense of *power*, in terms of access to the scarce resources; thus, whenever possible, people tend to prefer to belong to stronger groups. If one is unalterably assigned to a particular group, one tends to prefer to signal loyalty to this group and to increase its power. In this light, we note a third force: In questing for positive self-images, people seek approval from their peers or, alternatively, seek to avoid disapproval. This approval or disapproval is most powerful with physical presence and immediacy, diminishing when one's in-group members are 'out of sight, out of mind'.

In any case, it is not necessarily optimal to discriminate when playing a game, as the strategic environment means one must anticipate the choices of the other player(s). Our main hypothesis is the following: The presence of a salient group can affect the choice of individuals by making them proportionally more aggressive, for the reasons mentioned above. The more salient the group, the more the individuals in the group behave aggressively in the interaction with members of opposite groups. This aggressive behavior will lead to different outcome patterns in different strategic environments, impacting performance (social payoffs) in different ways.[5]

3. EXPERIMENTAL DESIGN & IMPLEMENTATION

Our experiment differs in two major respects from earlier experiments by social psychologists: First, we set up a strategic environment, and second, our group is not minimal. In our experiment, the group is made significant for the individual because the payoff of each member depends in part on the actions of the members of the group. But the payoffs do not change across different conditions: Whether the group is salient or not, the payoffs are determined in the same way. We examine whether the payoff-irrelevant changes concerning how the group witnesses the actions of the member affects the member's behavior, and how any such changes vary across different games.

[5] On the issue of performance, the social-facilitation literature documents the effect that the mere presence of an audience can have on human and animal *individual* performance. Zajonc, Heingartner, and Herman (1969) find that the performance of cockroaches in a simple task (finding food in a straight maze) is facilitated by the presence of an audience of cockroaches, while performance in a complex task (finding food in a maze with several turns) was impaired by that same type of audience. Of course, our setup is rather different, as each of our participants is a player in a game.

Design

In all treatments a stage game was played over multiple periods, with two action choices available for each of the two players in the stage game. Participants were randomly assigned to one of two groups: the Row group or the Column group. This assignment did not change throughout the experiment. Each group was taken to a separate room, labeled Room R and Room C; these labels were clearly written on the blackboard of the respective rooms. Therefore, we can think of a Row player as a "Host" in Room R and of a Column player as a "Guest" in that same room (and *vice versa* for Room C). In each room, play took place at a table with two seats on opposite sides, one for the Host player and one for the Guest player. Two cards, one for each of the possible actions, were placed face down on the table in front of each seat. Each participant played once in Room R and once in Room C, passing his or her choice face down across the table to the experimenter. Therefore, during a session, each player was a Host player once and a Guest player once.

In each stage game there were two active participants (the ones making choices) and the remaining participants were inactive. The payoff to each active participant was the sum of the payoffs from the outcomes in the two games in which he or she made choices. Inactive players (the ones who do not make a choice) also had a material interest in a game's outcome; each inactive player received a share of 1/3 of the outcome that corresponds to his or her group (Row or Column). We aimed for 20 participants and so 10 periods in a session, but sometimes in practice had only 16 or 18 people show up, so that there were only eight or nine periods in these respective cases. Thus, a large share of one's earnings was the result of choices made by the other members of one's group. All of this is common to all treatments.

Treatments differed according to whether or not there was an *audience* present when the Host and Guest players made their choices, and also differed according to whether or not we provided immediate *feedback* about the outcome of the stage game. Therefore, we have a classic 2x2 design. In the two treatments corresponding to the audience condition, some inactive players are present in the room when play takes place. These are the inactive Row players in Room R and the inactive Column players in Room C. The audience is therefore comprised of individuals who have a common interest with the Host player. These individuals were seated in a semi-circle behind the Host player and watched play unfold; the audience was not permitted to make sounds or gestures. In the two treatments corresponding to the no-audience condition, only the two active players and the experimenter were present in the room. The inactive Row and Column players waited in two separate holding rooms. In the feedback condition, the experimenter immediately exposed the cards chosen to everyone in the room. In the no-feedback condition, the experimenter receives the two cards and says nothing. So in the two treatments corresponding to this condition, no player knew the outcome of the play at any stage (except, partially, at the moment of the final payment). In the case of feedback with no audience, an experimenter accompanied the home participant back to the waiting room and announced the actions chosen.

To summarize, the timing is the following: (1) each participant is assigned the role of Row or Column player, and players are divided into two rooms; (2) instructions are read aloud (an example is provided in Appendix A) and questions answered; (3) play takes place; and (4) payments are made individually and anonymously. We had four conditions: AF (Audience and feedback), NAF (No audience and Feedback), together with ANF and NANF. Cash payments were made at the end of the experiment, individually and privately.

Implementation

As mentioned, we chose two 2x2 games, the Battle of the Sexes and the Prisoner's Dilemma. These are shown below in Figures 1 and 2:

Figure 1: Battle of the Sexes

	A	B
A	3, 1	0, 0
B	0, 0	1, 3

Figure 2: Prisoner's Dilemma

	A	B
A	5, 5	1, 7
B	7, 1	2, 2

The Battle of the Sexes has two pure-strategy equilibria, (A,A) and (B,B), as well as a mixed-strategy equilibrium in which each player choose the preferred action with probability 0.75. In this equilibrium the coordination rate (probability of the non-zero outcome) is 0.375. Our Prisoner's Dilemma is standard; the specific numerical values are chosen so that the joint value of cooperation is larger than the joint value of defection and cooperation. This makes the choice of cooperation more appealing.

Both games have been extensively studied experimentally. Cooper, DeJong, Forsythe and Ross (1989) consider the Battle of the Sexes, with the same parameter specification (up to a multiplicative factor) as ours; their experiment consisted of a sequence of one-shot games.[6] Their evidence suggests that subjects converge to a frequency of choices that is close to the

[6] The subjects were matched in a sequence of 20 one shot games. Each game was played with a different partner, and subjects knew this. The payoffs were expressed in terms of a lottery with a single prize of one dollar.

mixed-strategy equilibrium, with a smaller proportion on the favorite action than the Nash proportion, 63% rather than 75%. In an extension of the experiment, players were allowed to make simultaneous independent announcements of their intended choice. When the two announcements resulted in a coordination of a non-zero outcome, the decisions followed the announcement 80% of the times. When the announcements did not match, the average proportion of choice of the favorite action for each player increased to a 71%. Similar results are in Cooper *et alii* (1990, 1992, 1993) and Straub (1995).

A multitude of experiments have been conducted on the prisoners' dilemma (see Rapaport and Chammah 1965, Dawes 1980, and Roth 1988 for surveys of these experiments). While it is well-known that people cooperate (choose the dominated action A) with positive frequency in one-shot matches (typically 10-30%), a central finding in these studies is that mutual cooperation is indeed rather rare in the prisoner's dilemma. We chose a calibration that seemed likely to attract a fair amount of cooperation, so that we could more easily identify effects across audience and feedback treatments.

We had three sessions for each of the four conditions in each game. Thus, we conducted 12 sessions for each game, or 24 sessions in all. Participants were students at the University of California at Berkeley or the University of California at Santa Barbara, recruited from the general campus population. On average, people received about $18 (including a $5 show-up fee) for a session lasting about an hour. A total of 228 subjects participated in the Battle of the Sexes experiment, each playing twice; 112 games were played with an audience, and 56 of these had feedback. Of the 116 games played without an audience, 56 had with feedback. A total of 206 subjects participated in the Prisoner's Dilemma experiment; 100 games were played with an

audience, and 50 of these had feedback. Of the 106 games played without an audience, 58 had feedback.

4. RESULTS

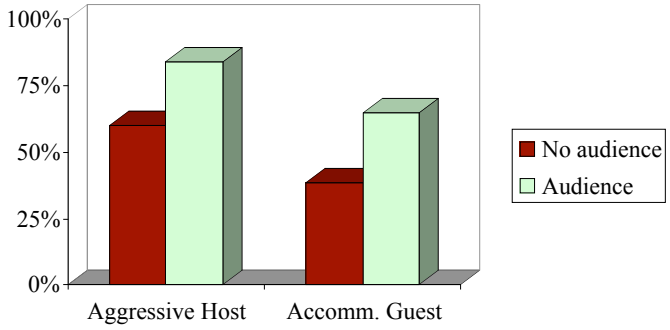
The presence of an audience affects players' behavior and the game's outcome in the two games we study. This effect is present whether or not the outcome of play is immediately revealed, and is substantially stronger when this feedback is given. We can conveniently summarize more detailed results by pooling data for Host and Guest choices.[7] Our experiments show that when an audience is present (and feedback is provided) Host players tend to make choices that favor the audience, while Guest players largely anticipate this behavior.

Battle of the Sexes

The two actions have a simple intuitive meaning in this game. For the Row player, action A aims at the largest payoff for oneself while B aims at largest payoff for the other player. For the Column player, B aims at the largest payoff for oneself while A aims at largest payoff for the other player. Thus, we say the Host player is *aggressive* when choosing an action aiming at the largest payoff for his or her role (A for Row and B for Column). We say the Host player is *accommodating* otherwise. Similarly, the Guest player is *aggressive* when choosing an action aiming at the largest payoff for his or her role and *accommodating* otherwise. This is of course simply a convenient classification. In Figure 3, we show how the frequency of aggressive or accommodating choices is affected by whether an audience is present:

[7] Recall that Host players are the Row players in Room R and the Column players in Room C, while Guest players are the Column players in Room R and the Row players in Room C.

Figure 3 - Audience Effects - Battle of the Sexes



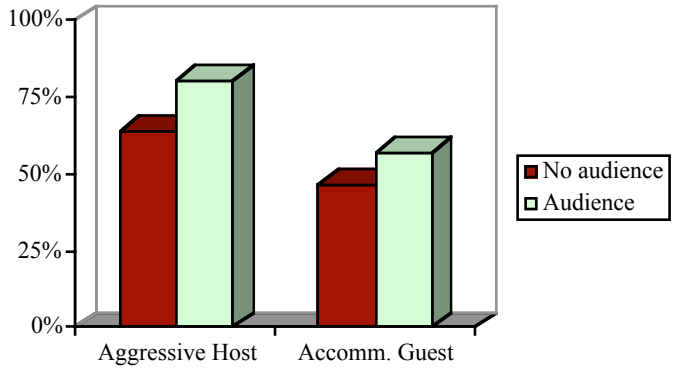
The presence of an audience increased the probability of the Host making the aggressive choice increased from 59.8% (67 of 112 observations) to 83.6% (97 of 116 observations). Similarly, the probability of the Guest making the accommodating choice increased from 38.4% (43 of 112 observations) to 64.7% (75 of 116 observations). Thus the presence of an audience shifted behavior by about 25 percentage points for each of the Host and the Guest.

The effect is also statistically significant. For example, a Wilcoxon-Mann-Whitney rank sum test based on every observation rejects the hypothesis that choices have the same distribution with and without an audience: For the Host ($Z = 3.99$, $p < 0.0001$) as well as for the Guest ($Z = 3.96$, $p < 0.0001$). Of course these observations may not be independent, since the previous history of actions may affect later behavior. Nevertheless, the effect is still significant for both Hosts ($Z = 3.20$, $p = 0.001$, two-tailed test) and Guests ($Z = 2.20$, $p = 0.028$, two-tailed test) when we consider only observations from the first period, where no interaction takes place. This confirms our main finding: There is a significant difference between behavior in front of an interested partisan audience and behavior when there is no audience.

In our experiment, the feedback condition involves more than just the simple provision of information. When an audience is watching, the provision of feedback means that all the

observers know what a player had just chosen, so to the informational consequences we must add the psychological impact. Figure 4 shows how the frequency of aggressive or accommodating choices (pooling R and C players) is affected by whether feedback is given:

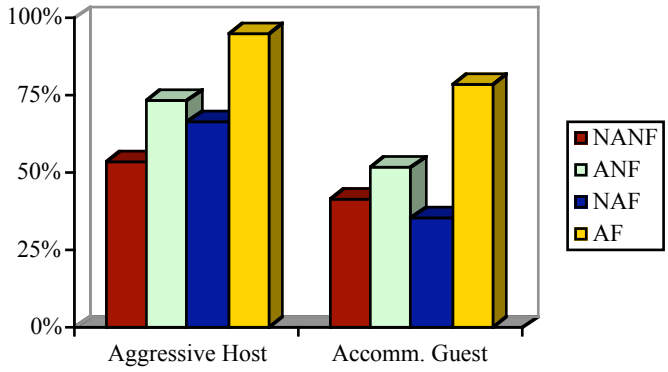
Figure 4 - Feedback Effects - Battle of the Sexes



The provision of feedback increased the probability of the Host making the aggressive choice increased from 63.8% (74 of 116 observations) to 80.4% (90 of 112 observations). Similarly, the probability of the Guest making the accommodating choice increased from 46.6% (54 of 116 observations) to 57.1% (64 of 112 observations). The direction of the effect is the same as for audience presence, but the difference is smaller. It is statistically significant if we consider the entire data set in the case of hosts (Z=-2.78, p = 0.005 in the Wilcoxon-Mann-Whitney rank sum test), but it is not for guests (Z= -1.59, p = 0.11) even when we treat each observation as being independent.

We have so far considered the effect of the two treatments, audience and feedback, separately; however, as we shall see, there is also a strong interaction effect. Figure 5 illustrates the effects on Host and Guest behavior, disaggregated according to combinations of audience and feedback:

Figure 5 - BoS - Play by Audience and Feedback



Hosts played aggressively 53.6%, 73.3%, 66.1%, and 94.6% of the time in the NANF, ANF, NAF, and AF treatments, respectively; Guests made accommodating choices 41.1%, 51.7%, 35.7%, and 78.6% of the time in these treatments. The most noteworthy feature of these data is that there is a very powerful effect when the presence of a partisan audience and the provision of feedback are combined; nearly all of the Hosts are aggressive with this highly-salient form of observation, and the preponderance of Guests correctly anticipate this behavior.

We can test for interaction effects by keeping one condition constant and varying the other. When feedback is provided, the Wilcoxon-Mann-Whitney test rejects the null hypothesis that the distribution of choices is the same with and without an audience for both Hosts (Z = 3.79, $p < 0.001$) and for the Guests (Z = 4.56, $p < 0.001$). However, the result is slightly different in the no-feedback condition. While the same null hypothesis is again rejected for the Hosts (Z = 2.20, $p < 0.028$), we cannot reject it for the Guests (Z = 1.14, p = 0.255). It appears that audience and feedback need to both be present for significant effects on Guest behavior.

We also find that the effect of feedback provision on behavior is significant, but only when the audience is present. In this case, the Wilcoxon-Mann-Whitney test rejects the null hypothesis that the distribution is the same with and without feedback, both for Hosts (Z = 3.09,

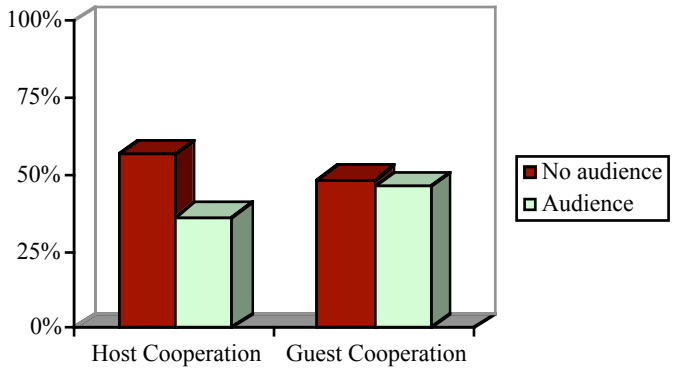
p = 0.002) and for Guests (Z = 3.02, p = 0.003). However, this null hypothesis cannot be rejected when no audience was present, for either Hosts (Z = 1.34, p = 0.179) or Guests (Z = 0.580, p = 0.562).[8]

In the Battle of the Sexes, we therefore conclude that the audience effect is large even when no feedback is provided and providing feedback makes a difference if and only if the audience is present.

Prisoner's Dilemma

As is usually done in the interpretation of the Prisoner's Dilemma, we describe the choice of action A as cooperation and that of B as defection. In Figure 6, we show how the frequency of cooperative choices (pooling R and C players) is affected by whether an audience is present:

Figure 6 - Audience Effects- Prisoner's Dilemma



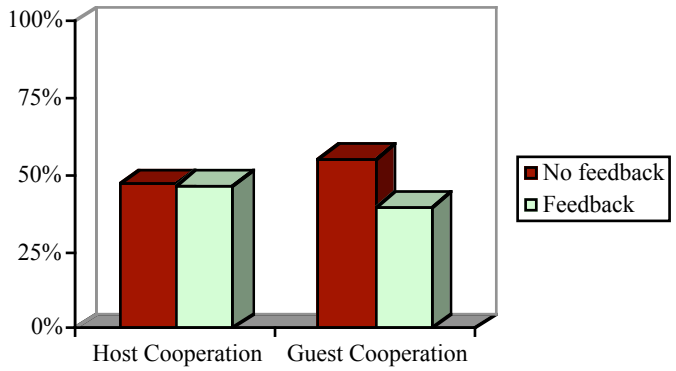
As we found in the Battle of the Sexes, the audience effect is noticeably different for Hosts and Guests. The presence of an audience decreased the probability of Host cooperation from 56.7% (60 of 106 observations) to 36.0% (36 of 100 observations); this difference is

[8] We can obtain a more synthetic analysis of the observed effects using logit regressions. In the Battle of the Sexes, a regression of Host choice of aggressive play on audience, feedback and the interaction of audience and feedback gives a significant coefficient for audience (Z = 2.19, p = 0.028, marginal effect 0.15) and a marginally-significant coefficient for the interaction of audience and feedback (Z = 1.74, p = 0.082, marginal effect 0.19).

statistically significant (Z = 2.96, p = 0.003). On the other hand, the probability of Guest cooperation was roughly the same with or without an audience, 48.1% (51 of 106 observations) or 46.0% (46 of 100 observations), respectively; this difference is not significant (Z = 0.30, p = 0.762). Overall, the aggregate cooperation rate is 52.4% with no audience and 41.0% with an audience, and this difference is statistically significant (Z = 2.31, p = 0.021). If we only consider the first-period data, the decrease in Host cooperation is similar, going from 58.3 to 33.3%. However, in contrast to the case of the Battle-of-the-Sexes game considered earlier, here the audience effect is not significant (Z = 1.20, p = 0.22) with this limited number of observations.

We next consider the effect of feedback in the Prisoner's Dilemma, aggregating across audience conditions, as shown in Figure 7:

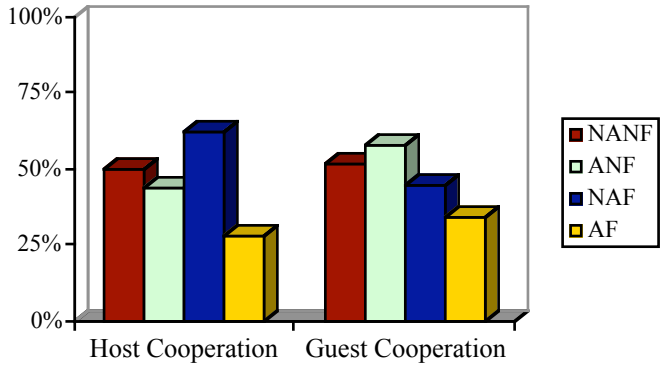
Figure 7 - Feedback Effects- Prisoner's Dilemma



In the aggregate, the provision of feedback had no effect on Host cooperation, which was 46.9% (46 of 98 observations) with feedback and 46.3% (90 of 112 observations) without feedback; naturally this small difference is not significant (Z = 0.092, p = 0.927). However, feedback does appear to reduce Guest cooperation rates, from 55.1% (54 of 98 observations) to 39.8% (43 of 108 observations); this difference is significant (Z = 2.19, p = 0.029).

We can also examine the interaction of audience and feedback in the Prisoner's Dilemma; Figure 8 illustrates the effects on Host and Guest behavior, disaggregated according to combinations of audience and feedback:

Figure 8 - PD - Play by Audience and Feedback



Hosts cooperated 50.0%, 44.0%, 62.1%, and 28.0% of the time in the NANF, ANF, NAF, and AF treatments, respectively; Guests cooperated 52.1%, 58.0%, 44.8%, and 34.0% of the time in these treatments. Once again, the strongest effect occurs when the presence of a partisan audience and the provision of feedback are combined, as cooperation is rather sharply reduced in this case.

Regarding interactions between audience and feedback, the presence of an audience has no significant effect on Hosts when no feedback is provided ($Z = 0.59$, $p = 0.554$); however, the presence of an audience induces a large (34 percentage points) and significant ($Z = 3.52$, $p < 0.001$) reduction in the Host cooperation rate when feedback is provided. For Guests, there is no significant effect either without feedback ($Z = -0.59$, $p = 0.558$) or with feedback ($Z = 1.14$, $p = 0.254$).

Providing feedback increases Host cooperation when there is no audience, but the change is not significant ($Z = -1.24$, $p = 0.214$); however, feedback has the opposite effect on Hosts

when there is an audience watching, with this change marginally-significant ($Z = 1.66$, $p = 0.097$, two-tailed test). The pattern for Guests is rather similar: Providing feedback has no significant effect on cooperation when there is no audience ($Z = 0.74$, $p = 0.459$), but reduces cooperation significantly ($Z = 2.40$, $p = 0.017$) when an audience is present.[9]

In the Prisoner's Dilemma, we therefore conclude that there is a strong interaction between audience and feedback in the choice of Hosts: the audience effect is stronger when combined with feedback.

Social efficiency

A fundamental and extremely controversial issue in the analysis of group behavior in social sciences is whether group behavior affects the efficiency of the outcomes. Here we consider *social efficiency*, by which we simply mean the total of the payoffs generated; of course social efficiency and individual self-interest may suggest very different choices. It is important and instructive that salient group membership has quite different effects on social efficiency in our two games, as we see substantially more coordination in the Battle of the Sexes when there is both an audience and feedback, and considerably less cooperation in the Prisoner's Dilemma under these conditions.

In Table 1 we report the rate of successful coordination for our treatments of the Battle of the Sexes. The frequency of each of the four outcomes by treatment is reported in detail in Appendix B.

[9] The only significant coefficient in the logit regression parallel to the one we described for the Battle of the Sexes is that for the interaction term ($Z = -2.17$, $p = 0.039$, with marginal effect -0.276).

Table 1: Successful coordination in the Battle of the Sexes

	NA NF	NAF	ANF	AF
Coordination rate	58.9% (33 of 56)	55.4% (31 of 56)	55.0% (33 of 60)	73.2% (41 of 56)
Proportion favoring Hosts	.455 (15 of 33)	.516 (16 of 31)	.727 (24 of 33)	1.000 (41 of 41)

In every treatment, the rate of success was higher than the rate for the mixed-strategy Nash equilibrium. We note that the success rate in the AF treatment was substantially higher than in the other treatments, which all have similar rates of successful coordination. When we consider the proportion of successful outcomes that favored the Host, the presence of an audience is seen to be a very strong force: Pooling the two no-audience treatments, the overall proportion favoring the Host is 48.4% (31 of 64), very nearly the one-half that would be expected with random behavior. In contrast, 87.8% (65 of 74) of all successful outcomes favored the Host when an audience was present. In fact, when there was both an audience and feedback, every one of the 41 successful outcomes favored the Host. In general, when there was an audience players strongly tended to choose their favorite action when playing as Hosts and tended to concede when they play as Guests.

In Table 2 we report the rate of mutual cooperation and social efficiency for our treatments of the Battle of the Sexes. The frequency of each of the four outcomes by treatment is reported in detail in Appendix B. Social efficiency is defined as the extent to which the average joint payoff exceeded the minimum (four, with mutual defection) divided by the amount that this joint payoff could possibly exceed this minimum (six, with mutual cooperation). We also show the proportion of times that an off-diagonal outcome favored the Host.

Table 2: Cooperation & Social Efficiency in the Prisoner's Dilemma

	NANF	NAF	ANF	AF
Rate of mutual cooperation	.208	.293	.260	.060
Rate of mutual defection	.187	.224	.240	.440
Social efficiency	.611	.615	.593	.393
Proportion of off-diagonal outcomes favoring Hosts	.517	.321	.640	.560

The joint effect of audience and feedback is clear, as mutual defection has a frequency of .440, compared to a frequency between .187 and .240 in the other three treatments. A complementary result is that the frequency of mutual cooperation falls to .060 from a range between .208 and .293 in the other three treatments. There is a corresponding drop in social efficiency, the level of which is quite similar in the other three treatments. We also see that the off-diagonal outcomes favor the Host more when there is an audience (.640 and .560, compared to .517 and .321) and when there is no feedback (.517 and .640, compared to .321 and .560).

5. CONCLUSION

We examine the effect of group membership on individual behavior in a precise, specific experimental setup. The results confirm the hypothesis that the saliency of one's group membership affects one's behavior, even if this saliency provides no information and has no effect on payoffs. Overall, this provides support for the original intuition in social psychology

that people automatically look, perhaps implicitly, to group membership as a guide for conduct in social situations.

The most significant and robust result, common to both games, is that the presence of a partisan audience and the provision of feedback together cause the Hosts move in the direction of their most preferred action, namely the action that gives the highest payoff. This is in agreement with the prediction that the presence of an audience makes the perception of power to the Host group stronger, and affects their expectations or sense of entitlement. Guests appear to perceive this effect on Hosts, even before they have the opportunity to observe the actions they make. For example, we have seen that in the Battle of the Sexes the Guests concede significantly more when an audience is present than they do when it is not, even in the first round, before they have seen any choice of any other subject in the experiment. This is an effect on behavior that is apparently produced through introspection.

In the Battle of the Sexes the salience of power pushes the players to coordinate on the outcome with payoff (3,1) or (1,3), depending on the location. The result is an increase of the average payoff compared to the no-audience environment. In the Prisoner's Dilemma Game the aggressive stance of the Hosts produces a shift toward the unique Nash equilibrium outcome where both players defect and the result is a reduction of the average payoff. However, power by itself is an insufficient explanation for the effect of immediate feedback on behavior when an audience is present. When feedback is provided, the host player may wish to seek the approval or to avoid the disapproval of his or her fellow in-group members, and therefore make the expected aggressive play. Informed and approving or disapproving audience members are literally in one's face, while the non-audience claimants are not.

This effect does not necessarily produce more conflict. It does in the Prisoner's Dilemma, but not in the Battle of the Sexes. Perhaps the reason for the difference is that in the latter game an implicit agreement is reached that people who do act aggressively as Guests will be rewarded when they play as Hosts. A similar agreement might be reached in the Prisoner's Dilemma game, with an appropriate choice of payoff values. For instance, one might choose a payoff for the outcome (A, B) that is less unequal than the (7, 1) in our sessions, or an even more appealing total payoff for this outcome. On the other hand, it may be the case that salient group membership tends to reinforce equilibrium play, so that benefits can be obtained only when there are multiple equilibria in the game at hand.

The behavior of our subjects in the Prisoner's Dilemma experiment can be labeled self-defeating group solidarity. There are examples in history where such solidarity has lead to disaster. Diamond (2004) describes the catastrophic end of the Norse colony in Greenland. He emphasizes as the reason for the collapse the inability to manage a fragile environment. At the root of this inability however was the unwillingness to accept technology, lifestyle and even diet close to those of the indigenous population, the Inuit. McGovern (1994) also argues that for the Norse, ethnic purity triumphed at the expense of biological survival. While the starving settlers slaughtered their cattle and dogs, "there were seals in the fjord, right under the ice". But without harpoons and the skill to find the seals' breathing holes in the ice, the Norse couldn't reach them. Those customs were more suitable for that environment, but unacceptable for the Norse because they would threaten their social identity.

It seems clear that group influence and social identity are important issues that can be ignored only at the peril of researchers in the social sciences. Many extensions of our work seem immediate: These include testing behavior when the audience is smaller, the audience is mixed,

the audience does not share in the payoffs, or feedback is uncertain. Our first study should encourage other researchers to investigate how membership in groups affects the behavior of individuals in strategic environments and economic institutions.

APPENDIX A

INSTRUCTIONS (room R)

Thank you for participating in this experiment. You will receive $5 for your participation, in addition to other money to be paid as a result of decisions made in the experiment.

There are 20 people participating in this session. They have been randomly divided into two rooms, each with 10 people. You are in room **R**, this means you are a **Row decider**.

There will be ten rounds in this session, and each person will make two decisions, one in each room. You have a card with a green number and a card with a (different) yellow number. These numbers will determine when and where you make decisions.

Your **green number** indicates the round during which it will be your turn to make a decision in the **room where you are now (room R)**.

Your **yellow number** indicates the round during which it will be your turn to **go to the other room (room C)** and make a decision there.

In each round there are two people making a decision. Each person will be making a simultaneous choice between A and B in the following decision matrix:

		Column	
		A	*B*
Row	*A*	5 , 5	1 , 7
	B	7 , 1	2 , 2

In each cell, the first number represents the outcome for the Row decider and the second number represents the outcome for the Column decider.

Thus, if both people choose A, the Row decider receives 5 and the Column decider receives 5. If both people choose B, the Row decider receives 2 and the Column decider receives 2. If the Row decider chooses A and the Column decider chooses B, the Row decider receives 1 and the Column decider receives 7. If the Row decider chooses B and the Column decider chooses A, the Row decider receives 7 and the Column decider receives 1.

The other nine members of each room also have a financial stake in the outcome – each person not making a decision receives 1/3 of the amount shown for the realized outcome.

Thus, if both deciders choose A, every inactive person in room R receives 5/3 and every inactive person from room C receives 2. If both deciders choose B, every inactive person from room R receives 2/3 and every inactive person from room C receives 2/3. If the Row decider chooses A and the Column decider chooses B, every inactive person from room R receives 1/3 and every inactive person from room R receives 7/3. If the Row decider chooses B and the Column decider chooses A, every inactive person from room R receives 7/3 and every inactive person from room R receives 1/3.

Each unit is worth $0.50 in actual money (2 units = $1) that will be paid in cash at the end of the experiment.

All people in the room (except for the person from the other room) will be able to watch the decider who belongs to their room make his or her choice (however, no verbal comments are permitted).

The decision of the person who walks into the room, on the other hand, is made privately.

The outcome of the joint decision is immediately revealed to all people in the room.

After the 10 rounds are completed, we will total each person's earnings (from the outcomes of the two self-made decisions, as well as the other 18 outcomes), add the $5 show-up fee, and pay each person individually and privately, using the numbers on your two cards to identify your decisions.

Please feel free to ask questions.

APPENDIX B

Table B1: Frequency of outcomes in the Battle of the Sexes, by audience and feedback

No Audience No Feedback	Guest Aggressive	Guest Accommodating	**Audience No Feedback**	Guest Aggressive	Guest Accommodating
Host Aggressive	.268 (15)	.268 (15)	Host Aggressive	.333 (20)	.400 (24)
Host Accommodating	.321 (18)	.143 (8)	Host Accommodating	.150 (9)	.117 (7)
No Audience Feedback	Guest Aggressive	Guest Accommodating	**Audience Feedback**	Guest Aggressive	Guest Accommodating
Host Aggressive	.375 (21)	.286 (16)	Host Aggressive	.214 (12)	.732 (41)
Host Accommodating	.268 (15)	.071 (4)	Host Accommodating	.000 (0)	.054 (3)

Table B2: Frequency of outcomes in the Prisoner's Dilemma, by audience and feedback

No Audience No Feedback	Guest Cooperating	Guest Defecting	**Audience No Feedback**	Guest Cooperating	Guest Defecting
Host Cooperating	.208 (10)	.292 (14)	Host Cooperating	.260 (13)	.180 (9)
Host Defecting	.313 (15)	.187 (9)	Host Defecting	.320 (16)	.240 (12)
No Audience Feedback	Guest Cooperating	Guest Defecting	**Audience Feedback**	Guest Cooperating	Guest Defecting
Host Cooperating	.293 (17)	.328 (19)	Host Cooperating	.060 (3)	.220 (11)
Host Defecting	.155 (9)	.224 (13)	Host Defecting	.280 (14)	.440 (22)

REFERENCES

Billig, Michael and Henri Tajfel (1973), "Social categorization and similarity in intergroup behavior,'' *European Journal of Social Psychology*, **3**, 27-52.

Cooper, Russell, Douglas DeJong, Robert Forsythe, and Thomas Ross (1989), "Communication in the Battle of the Sexes: Some experimental results,'' *Rand Journal of Economics*, **20**, 568-587.

Cooper, Russell, Douglas DeJong, Robert Forsythe, and Thomas Ross, "Selection criteria in coordination games: Some experimental results," *American Economic Review*, **80**, 218-233, (1990).

Cooper, Russell, Douglas DeJong, Robert Forsythe, and Thomas Ross (1992), "Communication in coordination games," *Quarterly Journal of Economics*, **107**, 739-771.

Cooper, Russell, Douglas DeJong, Robert Forsythe, and Thomas Ross (1993), "Forward induction in Battle of the Sexes games," *American Economic Review*, **83**, 1303-1316.

Diamond, Jared (2004), *Collapse: How societies choose to fail or succeed*, Viking.

Huntington, Samuel (1996), *The clash of civilizations and the remaking of the world order*, Simon and Schuster.

Mayo, Elton (1949), *The social problems of an industrial civilization*, London, Routledge and Kegan Paul.

McGovern, Thomas (1994), "Management for Extinction in Norse Greenland," in C. Crumley, ed., *Historical Ecology: Cultural Knowledge and Changing Landscapes.* Santa Fe: School of American Research Monograph, 127-154.

Straub, Paul (1995), "Risk dominance and coordination failure in static games," *The Quarterly Review of Economics and Finance*, **35**, 339-363.

Tajfel, Henri, Michael Billig, R. P. Bundy, and Claude Flament (1971), "Social categorization and intergroup behavior,'' *European Journal of Social Psychology*, **1**, 149-178.

Tajfel, Henri and Turner, John (1979), "An integrative theory of intergroup conflict," in W. G. Austin, and S. Worchel, *The social psychology of intergroup relations*, Brooks Cole, Monterey, CA.

Taylor, Frederick W. (1911), *Principles of scientific management*, New York, Harper.

Zajonc, R., A. Heingartner, and E. Herman (1969), "Social Enhancement and Impairment of Performance in the Cockroach," *Journal of Personality and Social Psychology,* **13**, 83-92.