WORKING PAPER
DIPARTIMENTO DI ECONOMIA PUBBLICA

Working Paper n.183

Giuseppe Ciccarone, Giovanni Di Bartolomeo e Stefano Papa

Is social identity belief independent?

Roma, Luglio 2018

SAPIENZA
UNIVERSITÀ DI ROMA

Is social identity belief independent?*

Giuseppe Ciccarone, Giovanni Di Bartolomeo, Stefano Papa

Sapienza University of Rome

This version: 5 July 2018

Abstract:

In this paper we aim to disentangle the effects on in-group favoritism driven by beliefs from those stemming from group identity, with the final goal of testing the relative power of three potential explanations of this bias: The *Beliefs Driven Explanation* (*BDE*), the *Group Identity Explanation* (*GIE*) and the *Belief-mediated Group Identity Explanation* (*BGE*). The *BDE* suggests that in-group favoritism is only driven by the desire not to let others' expectations down. The *GIE* claims that people have a preference, *per se*, for members of their group. According to the *BGE*, people also have a preference for members of their group, but this is mediated by their second-order beliefs. To this aim, we built an experimental design able to produce exogenous variations in both group membership and expectations, hence providing a genuine test for the rationale of in-group bias. The results of our experiment suggest that beliefs *per se* are not a significant explanation of in-group favoritism and hence do not provide support to the *BDE*. Our experimental evidence does not provide support also to the *BGE*. We conclude that our experiment suggests to single out the GIE as the most powerful explanation of social identity.

***JEL* codes:** A13, C91, D03, D64.

Keywords: Social identity, second-order beliefs, guilt aversion, causation.

* The authors are grateful to an anonymous referee, Maria Bigoni and John Hey for useful comments on previous drafts. Financial support by Sapienza University of Rome is gratefully acknowledged.

1. Introduction

Social Identity Theory (SIT) predicts that intergroup behavior depends on perceived group status differences, the perceived legitimacy and stability of these differences, and the perceived ability to move from one group to another (Tajfel and Turner, 1979, 1986). In-group favoritism is a common feature of such behavior, as individuals often endorse resource distributions that would maximize the positive distinctiveness of an in-group, in contrast to an outgroup, at the expense of personal self-interest. According to SIT, even when group identity is built on random basis, individuals belonging to a group tend to favor individuals belonging to the same group. Several contributions to the recent experimental literature exploring why in-group favoritism emerges confirm that social identity fosters trust and cooperation.[1]

The aim of this paper is to contribute to this field of research on social identity by providing evidence on the possibility to disentangle the effects on in-group favoritism driven by beliefs from those stemming from group identity. The motivation for this attempt is to move a step forward in the understanding of the relative explanatory power of potential explanations of the influence of social identity on cooperation. The main of these explanations, which differ for the weights attached to beliefs and to group identity, can be grouped as follows.

1. People dislike to let the others' expectations down, independently of (present) group membership. Hence, an in-group bias effect emerges if a group member believes that the other members have "higher expectations" about the members' behavior, i.e., if they expect their group members to be more cooperative with individuals belonging to their group than with people belonging to other groups (*Beliefs Driven Explanation*, *BDE*).[2]
2. People show a bias for in-group favoritism because they have a preference for group members *per se*. The in-group bias effect is therefore due only to individual preferences and it emerges independently of beliefs (*Group Identity Explanation*, *GIE*).[3]
3. Beliefs and group identity are not mutually exclusive and can hence interact, as people dislike to let down only the expectations of the individuals belonging to their group (who they care for) while not caring about the expectations of the members of other groups. A group member with a

[1] See, among others, Chen and Li (2009), Guth *et al*. (2009), Chen and Chen (2011), Guala *et al*. (2013), Ockenfels and Werner (2014). Section 2 briefly discusses the economic literature on social identity.

[2] This idea is mainly connected to the concept of guilt aversion popularized for promise keeping by Charness and Dufwenberg (2006). See also Battigalli and Dufwenberg (2007); Dufwenberg and Gneezy (2000); Guerra and Zizzo (2004); Bacharach *et al*. (2007); Reuben *et al*. (2009); Chang *et al*. (2011); Bellemare *et al*. (2011); Bracht and Regner, (2013); Regner and Harth (2014); Khalmetski *et al*. (2015).

[3] The idea of social norms as moral commitment is discussed in, e.g., Braver (1995), Ostrom *et al*. (1992), Ellingsen and Johannesson (2004), Vanberg (2008) and Ellingsen *et al*. (2010).

preference for the other group members who does not want to let their belief down will hence favor them *vis à vis* the members of other groups. In this case, the in-group favoritism due to group identity is mediated by the members' beliefs (*Belief-mediated Group Identity Explanation, BGE*).[4]

The attempt to establish the relative explanatory power of these explanations is not trivial, as the empirical unraveling of the effects of beliefs and group identity is made difficult by an endogeneity problem. To provide a simple example, assume that we attempt to test *BDE* vs. *GIE* by eliciting expectations in an experiment. If we observe both in-group favoritism and second-order beliefs, we can be tempted to conclude that we have evidence in favor of *BDE*. However, as correlation is not causation, it might well be true that the social identity effect is instead driven by a group identity affecting beliefs. If the individuals' behavior is driven by their preference for group members, but people understand the effects of group identity, they will believe that for this reason group members will be more cooperative with other in-group members than with out-group individuals. When meeting with another member belonging to the same group, an individual will hence expect that the other one cultivates greater beliefs about their gains and, as a result, we will again observe in the experimental outcome in-group favoritism in both cooperation and second-order beliefs.

In order to overcome this difficulty, we built an experimental design able to produce exogenous variations in both group membership and expectations in a modified dictator's game, which can provide a genuine test for the rationale of the in-group bias. By producing exogenous variations in group membership and comparing the behavior of individuals within the same group to that of individual outside the same group, independently of their expectations, we can test if individuals suffer from in-group bias independently of their expectations. By producing exogenous variation in expectations, we can instead test if individuals within the same group (or outside the group) behave differently when they face different second-order beliefs. By combining both exogenous variations, our methodology allows us to explore *BDE*, *GIE* and *BGE* and hence to experimentally identify the most convincing explanation of social identity.

Our experiments add to the literature on social identity, expectations and belief manipulation. In particular, we elaborate on the exogenous mechanism of belief manipulation proposed by Ockenfels and Werner (2014). In one of their experiments (i.e., experiment 2), dictators can strategically manipulate the recipients' believes. Before the transfer occurs, they may choose to learn whether or not the opponent shares the dictators' own group identity. Two treatments are then considered. In one treatment, if the

[4] See Ederer and Stremitzer (2018).

dictator decides to be informed about the recipient's group identity, the same information will be given to the recipient. In the other treatment, the same information will not be shared with recipients. As the bias towards in-group matches turns out to be substantially smaller for informed dictators in the treatment where the recipient is not informed about the dictator's group affiliation, Ockenfels and Werner (2014) conclude that this evidence supports the view that in-group favoritism is partly belief-dependent. Their argument is based on the idea that when dictators decide to be informed about the recipient's group identity and this information is shared, recipients have higher expectations compared to the case when dictators decide to be informed about the recipient's group identity but this information is not shared. As a consequence, in the two treatments dictators have different second-order beliefs. The two treatments remain however different in another aspect, which could drive the results in the same way as differences in beliefs can do, as the choice of disclosing information is not random and can be motivated by different strategic considerations. This generates a self-selection problem which does not make the two sample fully comparable.[5]

As we will show, we aim instead to fully disentangle expectations from other elements which could affect in-group favoritism by using a random device. Our manipulation is hence subtler and more articulated than that proposed by Ockenfels and Werner (2014), which can be conceived as an "extreme" test of our hypothesis.[6]

The results of our experiments suggest that beliefs *per se* are not a significant explanation of in-group favoritism and do not hence fully support *BDE*. They suggest instead that group membership, and hence *GIE*, is significant, and that in-group favoritism due to group identity does not seem to be mediated by beliefs, as instead predicted by *BGE*.

The paper is organized as follows. Section 2 briefly discusses the related literature. Section 3 introduces the experimental design and the experiment procedures. Section 4 presents and discusses the result. Section 5 concludes.

2. Related literature

Our paper is clearly related to the pioneering study of Tajfel *et al.* (1971), who were the first to document that people reward more in-group than out-group members by seeking to maximize either the in-group outcomes, or the differences between in-group and out-group outcomes. In their experiments, Tajfel *et al.*

[5] It is worth noting that Ockenfels and Werner (2014) show that dictators use strategically information disclosure (i.e., they behave differently in the two treatments).

[6] We thank an anonymous referee for this suggestion.

(1971) showed that this form of intergroup discrimination is basic and emerges even under a minimal setting (minimal group paradigm). According to the minimal group paradigm, experiment participants: (i) are randomly assigned to groups; (ii) are not allowed to communicate; (iii) do not know members of their in-group and of the out-group; (iv) have no vested interest in serving their group.

In the field of economic theory,[7] Tajfel and coauthors' finding fueled interest in social identity and opened up a new field of research on intergroup bias and discrimination. From the theoretical point of view, Akerlof and Kranton (2000) incorporated the psychology and sociology of identity into an economic model of behavior. They suggested that identity can be associated with different social categories and expected people's behaviors, considered as norms or prescriptions, and that individuals can be conceived as endowed with a neoclassical utility function that is negatively affected by deviations from the prescription (i.e., the deviation from group identity causes disutility). They showed that the inclusion of identity substantively changes the conclusions of previous economic analysis.[8]

From the experimental point of view, group membership was manipulated and participants were randomly assigned to interact with other persons classified as either in-group or outgroup members. This was mostly done either by using the minimal group paradigm (e.g., Chen and Li, 2009; Güth *et al.*, 2009),[9] or relying on natural groups, be these religious, organizational, geographic, ethnic, etc. (Goette *et al.*, 2006; Whitt and Wilson, 2007).[10] By using several laboratory experiments, in an extensive study, Chen and Li (2009) showed the effects of group identity that cause an increase in the social preferences on in-group members.[11] They found that in-group members gift, reciprocate, forgive and maximize overall efficiency significantly more than out-group members.

Klor and Shayo (2010) studied the effects of social identity on preferences over redistribution, highlighting a trade-off between in-group favoritism and monetary payoffs. Experiment participants systematically deviate from monetary payoff maximization to favor their group when the monetary cost of doing so is not too high. Chen and Chen (2011) proposed a group-contingent social preference model and derived conditions under which social identity changes equilibrium selection in coordination games.

[7] For a review of the social psychology of identity, see e.g., Brown (1986), Tajfel *et. al*. (1971), Tajfel (2010) and Wetherell (1996). Social psychology experiments are also discussed in Brown (1986: 541–566) and Wetherell (1996: 203–216).

[8] It is worth mentioning that Akerlof and Kranton's model has more than one thousand citations in the IDEAS database (see https://ideas.repec.org).

[9] Charness *et al.* (2007) did not find any effect of group membership on individual behavior when using the minimal group paradigm, whereas with other treatments where groups are more salient, group membership significantly affects individual behavior.

[10] Chen and Li (2009) did not find statistically significant differences between the behavior of participants assigned to groups based on their true painting preferences and that of participants randomly assigned to groups.

[11] Charness and Gneezy (2009) and Babcock *et al.* (2015) measured peer effects by exploiting exogenous assignment to peer groups and explored the effect of incentive programs on changing behaviors. The former showed that financial incentives can potentially promote healthy behaviors such as exercise, weight loss, and smoking cessation. The latter provided evidence in favor of monetary team incentives. Those can induce effort at lower cost than through direct individual payment.

Focusing on a minimum-effort game, they found that social identity has some significant effects on the game outcomes, but only when groups are enhanced by allowing pre-play communication. Focusing either on dictator's game or on trust games, evidence for in-group favoritism is found, among others, also by Fershtman and Gneezy (2001), Ben-Ner *et al*. (2009), Whitt and Wilson (2007), Hargreaves *et al*. (2009), Liebe and Tutic (2010), and Ioannou *et al*. (2012).[12]

Closely related to our aim, some recent studies explored the role played by beliefs in in-group discrimination. Güth *et al.* (2009) provided evidence, based on the observed correlations between dictators' self-reported beliefs and in-group favoritism, which is consistent with guilt aversion. However, correlations between transfers and beliefs may be prone to false consensus effects (Ross *et al*., 1977) and provide no conclusive evidence for beliefs to causally affect behavior.[13] The results of Ockenfels and Werner (2014) supported instead the view that in-group favoritism is only partly belief-dependent. Their evidence also supported the view that shared group identity *per se* changes charity preferences: substantially less in-group favoritism is observed when dictators are informed that recipients are unaware of the shared group membership.[14]

A related strand of the literature addressed also methodological issues. Our research is in fact connected to those introducing exogenous variation in beliefs and exploring behavior based on switching or similar mechanisms. Charness and Dufwenberg (2006) found a correlation between second-order beliefs and promise keeping. Switching mechanisms were then introduced to evaluate the causal nexus linking second-order beliefs to promise keeping (e.g., Vanberg 2008; Ellingsen *et al.* 2010; Di Bartolomeo *et al.* 2018). In this paper, we contribute to this field of research by investigating both the correlation and the nature of the causal relation linking second-order beliefs to in-group discrimination.

3. Experimental design and procedures

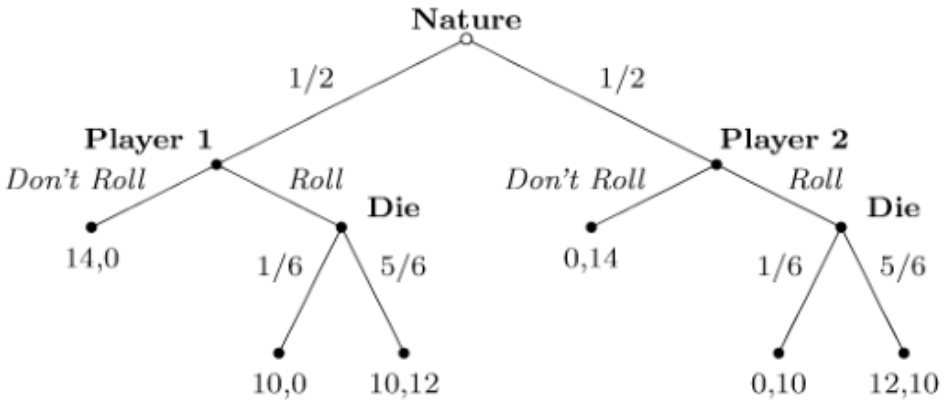
3.1 Experimental design and hypotheses

Our experiment is based on the binary-choice random-dictator game (Vanberg, 2008) described in Figure 1. In this game, *Nature* first randomly determines whether player 1 or 2 will be the *dictator*. The other player will be the *recipient*. The dictator chooses between two actions: *Don't Roll* and *Roll*. Choosing *Don't Roll*, the dictator receives 14 tokens and the recipient receives nothing; choosing *Roll*, the dictator

[12] Other studies are surveyed in Chen and Li (2009: Section I) and Balliet *et al*. (2014).
[13] See Vanberg (2008) for a discussion in the context of guilt aversion and promise keeping.
[14] See also Guala *et al.* (2013).

receives 10 tokens, whereas the recipient receives 12 tokens with probability 5/6 and nothings with probability 1/6.[15]

Figure 1. A binary-choice random-dictator game



Assigning all the players to two different groups in a random way (minimal group paradigm) at the beginning, the expected outcome of this game, which was actually obtained in our experiments,[16] is that the subjects' behavior is affected by group membership and that favoritism occurs.[17] If subjects are more likely to *Roll* when paired with a recipient belonging to the same group, we should observe in experiments a significant increase in both the rate at which dictators choose *Roll* and in their second-order beliefs, i.e., their beliefs about the recipients' beliefs that they will do so.

Such correlations are not however sufficient to fully understand the effects of social identity due to the relationship between actions and beliefs, as to this goal it is also necessary to investigate whether:

1. the increased frequency of *Roll* choices when participants are paired with recipients belonging to the same group is driven by changes in second-order beliefs induced by group membership (*BDE*), or by a preference for group membership;

[15] The payoffs and the names of choices in subgames match those used in the well-known Charness and Dufwenberg's (2006) trust game. Even though this setup is not strictly necessary to pursue the aim of the present paper, we here use Charness and Dufwenberg's (2006) setup because it guarantees that recipients cannot observe an egoistic action, i.e., *Don't Roll*, by the sender. As this experiment is part of a wider project carried out at CIMEO on effort determinants in asymmetric information contexts, in future research we plan to compare its results with those stemming from designs based on different mechanisms that could provide incentives for altruistic behavior (e.g., repeated interaction, promise giving, non-binding agreements, etc.).

[16] Results are reported in Section 4.1 and Appendix A.

[17] As discussed in Section 2, in-group favoritism has been documented in many similar contexts (see, e.g., Chen and Li, 2009; and Chen and Chen, 2010).

2. in the absence of evidence in favor of *BDE*, when in-group favoritism is explained by group identity, the effects of social identity are mediated by beliefs, that is, participants react (*BGE*), or do not react (*GIE*), to the different levels of beliefs induced by the experimenter.[18]

The simple comparison of "good matches" and "bad matches" – where by a good (bad) match we mean a dictator matched with a recipient of the same (different) color, i.e., an in- (out-) group match – is not helpful to understand the first issue, i.e., to disentangle between different explanations, as causation can run in both directions. The same comparison is also not helpful to understand the second issue, i.e., to understand the dictators' behavior for different second-order beliefs, as it could again suffer from endogeneity problems and the assignment in good or bad matches would be perfectly correlated with any associated change in beliefs.

In order to clarify the two above issues, we modify the random-dictator game by employing the switching mechanism introduced by Vanberg (2008) and, in this way, obtain independent variations in both group membership and second-order beliefs.[19] This modification, described below, is common knowledge for all participants.

In our experiment, all participants are first assigned to the "Red" team or to the "Blue" team according to the minimal group paradigm.[20] Pairs of subjects are then formed and the game is played. However, after Nature chooses the paired-subject roles and all dictators and recipients observe the color of their partners, some of the recipients are randomly re-matched with a new dictator, according to a given switching probability, which is known from the beginning. Each dictator whose co-player was switched observes the group color of the new recipient. After the re-matching, all recipients know that some of them could have been switched, but they do not observe whether they actually switched dictator. This way of creating exogenous variations in group membership introduces asymmetric information between dictators and recipients, as all subjects know they can be re-matched with a certain probability, but only dictators are told whether their recipient was switched, or not. A recipient's first-order beliefs (and hence a dictator's second-order beliefs) can therefore depend on whether she and her dictator are members of the same group and by the known switching probability, but not on whether they were effectively switched or not. After the re-matching, each dictator (be she involved or not in a switch) chooses between *Roll* and *Don't Roll*, and the payoffs are determined according to Figure 1.

The first explanation we test is *BDE*. If this explanation was true, and hence only beliefs mattered, the dictator's behavior would not depend on the actual switch, as membership after this shift is irrelevant. The

[18] Cf. Guth *et al*. (2009) and Ockenfeld and Werner (2014).
[19] The mechanism was introduced to study the rationale of promise keeping. In the same context, Di Bartolomeo *et al*. (2018) extended it to obtain an exogenous variation in expectations (second-order beliefs).
[20] This paradigm is based on minimal conditions to assign group memberships. See the next subsection for details.

recipient's first-order beliefs are instead formed before the shift and are affected by the initial match and by the switching probability (but are independent of the switch/no-switch condition). The dictator knows in fact that the recipient does not observe the shift and this cannot hence provide her with any information able to change the first order (the recipient's) beliefs. It follows that the second order (the dictator's) beliefs cannot change after the shift and no difference in the average roll rates between switch and no-switch dictators should be observe. Consider the case of a recipient who was initially matched with a dictator of the same color. If the switching probability is low, a dictator after the shift (independently of her initial group) believes that the recipient's beliefs are "high" and so she would be likely to roll not to disappoint these beliefs. If the switching probability is high, the dictator should believe that the recipient's beliefs are "lower" than before and should hence be less likely to roll.[21] If only second-order beliefs matter, *BDE* predicts that dictators should be more likely to roll whenever the partner was initially matched with a dictator of the same group, especially if the switching probability is low. Given the same initial match, the same behavior among actually switched and non-switched dictators should be observed.

If the results of the experiments did not confirm these predictions, the observed in-group bias would be consistent with a group identity explanation, but we must in this case further investigate whether group identity is mediated, or not, by beliefs. In order to properly disentangle *GIE* and *BGE*, we no longer need to produce an exogenous variation in group membership, but rather an exogenous variation in expectations, so as to compare the behavior of dictators matched with recipients of the same color holding however different second-order beliefs.

We produce this exogenous variation in expectations by using the extension of Vanberg's (2008) methodology introduced by Di Bartolomeo *et al*. (2018) in the context of promise keeping. This extension is based on the introduction of two different given and known switching probabilities (high and low). If group identity is consistent with in-group favoritism, then recipients' first-order beliefs, and dictators' second-order beliefs, will in fact depend on the value of the switching probability at the time dictators make their choices. Because of the existence of the group identity, we should expect non-switched dictators to be more inclined to roll when matched with a recipient belonging to their own group than with an outgroup member. Recipients who are initially matched with an in-group (out-group) member should expect the dictator to cooperate with higher (lower) probability if the switching probability is low rather than high. If dictators understand that, their second-order beliefs should vary with the switching probability and the initial group matching in the same direction as the recipients' first-order beliefs. This

[21] If a recipient was not matched with a dictator of the same color, the opposite should occur: if the switching probability is low (high), the dictator believes that the recipient's beliefs are "low" ("high") and so she would be less (more) likely to roll.

explains why our treatment variable, the switching probability, produces exogenous variations in first- and second-order beliefs.

We use that feature of the switching probability to test the effect of second-order beliefs in the in-group and in the out-group members:

i. a higher *Roll* rate among non-switched dictators with higher second-order expectations would imply that the dictators' behavior is causally affected by their expectations, as predicted by *BGE*;

ii. the same *Roll* rate, regardless of the second-order expectations, would imply that the dictators' behavior is not causally affected by their expectations, as predicted by *GIE*.

In the next session, we first test whether our data provide evidence in favor of in-group bias (the in-group *Roll* rates are higher than the out-group *Roll* rates) and of a positive correlation between *Roll* rates and dictators' second-order beliefs. As this evidence is confirmed by the experiment's results, in order to understand the causal nexus between beliefs and in-group favoritism, at this stage we introduce exogenous variations in both beliefs and membership. Our results show that the methodology we employ is actually able to produce exogenous variations in beliefs as: (i) the first order beliefs of in-group (out-groups) recipients under a low switching probability are greater (lower) than under a high switching probability; (ii) the dictators' second order beliefs reflect the recipients' first order beliefs.

Given this result, we test the different explanations of in-group bias.

H1: *Testing the BDE* (beliefs are the only determinant of in-group favoritism). Given the switching probability (which affects beliefs) and the recipient's initial match, we compare the behavior of switched dictators who are re-matched with out-group recipients with that of switched dictators who are re-matched with in-group recipients. If the dictators' behavior is not the same, then the *BDE* cannot hold.

H2: *Testing the GIE vs. BGE* (group identity is mediated by beliefs). In order to understand whether group identity is mediated by beliefs, as suggested by the *BGE*, we take group membership as given and compare the behavior of non-switched in-group dictators under a high and under a low switching probability (where beliefs will be shown to be different). Furthermore, we compare the behavior of switched in-group dictators re-matched with initially in-group recipients under high and low switching probability. Finally, we compare, under both switching probability, the behavior of in-group switched dictators re-matched with initially in-group recipients with that of in-group switched dictators re-matched with initially out-group recipients. If the dictators' behavior is the same in these circumstances, then the *BGE* does not hold and the *GIE* is the explanation with the greatest explanatory power.

Comparisons are made by using session averages and non-parametric test (Wilcoxon signed-rank test).

The robustness of our results is checked by using also a panel regression. These estimation outcomes are reported in Appendix D.

3.2 Procedures

The experiment was conducted in November 2017 at the *CIMEO Experimental Economics Lab* of Sapienza University of Rome.[22] The design involved 384 student subjects of this University (12 sessions, 32 subjects each), recruited using an online system. Upon arrival, subjects were randomly assigned to 32 isolated computer terminals.[23] Three assistants handed out instructions[24] and checked that participants correctly followed the procedures. Before playing any game, subjects filled out a short questionnaire testing their comprehension of the experiment.

The experiment consisted of three stages: group assignment (pre-session); experimental session; final payment (post session). Following Chen and Li (2009), we used the minimal group paradigm. Subjects were randomly assigned in two groups at the beginning of game, Red and Blue, and they remained in the same group until the end of the experiment (i.e., groups did not change throughout the experiment). All subjects were privately informed about their group membership and about the membership of the other participants.

After the group assignment, subjects played the experimental session, which consisted of ten rounds, with perfect stranger matching. Each round implemented the following sequence of five stages:

1. *Role assignment.* At the beginning of the round, roles were randomly assigned for each pair and subjects were informed of that. All dictators observed the *color* of the paired recipient and vice versa.
2. *Revelation of the switching probability*. Depending on treatment, the switching probability was announced (either 12.5% or 87.5%).
3. *Belief elicitation*. This stage has two parts:
    a. First-order beliefs: each recipient was asked to guess about his/her payoff.
    b. Second-order beliefs: each dictator was asked to guess the guess of the recipient with whom he/she had formed a pair.
4. *Switching*. Some pairs were switched, depending on the given switching probability. Only dictators were informed whether or not a switch occurred (recipients were not informed whether they had been switched or not). Dictators with switched recipients were also informed of: 1) the

[22] This is the acronym for *Center for Interpretation and Modeling of Experimental Observations.*
[23] The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).
[24] See the online Supplementary Material.

group's color of the new recipients; 2) the group's color of the previous partners (dictators) of their new recipients.

5. *Dictators' action*. All dictators made their choice: *Roll* or *Don't Roll*. All subjects were informed of their payoff for the round. Recipients could not infer the dictator's choice when their payoffs were zero.[25]

Finally, at the end of the session, subjects were paid. All subjects received a fixed show-up fee of 2.50 tokens. One of the rounds was randomly chosen for payment. The payoffs, as shown in Figure 1, were computed in tokens (where 1 token = 0.5 euro). At the end of each session, Incentives for beliefs elicitation were provided for all rounds, except the one chosen for payment, implying that subjects had no incentive to hedge against bad outcomes and thus to misreport their beliefs.[26]

It is worth stressing that, in line with Charness and Dufwenberg's (2006), we elicit beliefs before dictators make their choice, that is, before they know whether they have been switched or not. Exogenous variations in their expectations can hence be due only to a change in the switching probability, given the initial match. We assume that a change in the switching probability produces also exogenous variations in switched dictators' expectations.[27] This is reasonable because dictators know that recipients do not observe whether a switch occurred and so their first-order beliefs should depend on the switching probability and on the initial match, but not on the switch condition.[28]

4. Results

4.1 Expectations and switching probabilities

Before exploring the causal effect of in-group and out-group membership and of second-order beliefs, we must establish that our design is indeed able to induce exogenous variation in first- and in second-order beliefs. Table 1 presents recipients' (first-order) beliefs about dictator's decision to choose *Roll*. Columns refer to the cases of low and high switching probabilities; rows refer to matching: in-group refers to all the pairs where the dictator and the recipients were initially assigned to the same group color; out-group refers to the other pairs. Recipients are 1920; switch and no-switch subjects are pooled together (remember that recipients never know whether they are actually switched).

Table 1. Recipients' first-order beliefs (1920 obs.) *

[25] Recipients could obtain a zero payoff in two cases: (i) their dictator had chosen *Don't Roll*; (ii) their dictator had chosen *Roll* and the outcome of the die-roll was "1."

[26] Appendix B describes in detail how beliefs were elicited and computed.

[27] Recall that we elicit the beliefs of all dictators before the switch takes place, whereas we do not elicit the beliefs of *switched* dictators regarding the beliefs of their *re-matched* recipients.

[28] See Di Bartolomeo *et al*. (2018) for a more detailed discussion on the elicitation timing in the procedure here employed.

	Switching Probability	
	Low (12.5%) (1)	High (87.5%) (2)
Matching groups		
(A) Out Group	0.36 (s.d. 0.32, obs. 384)	0.45 (s.d. 0.31, obs. 528)
(B) In Group	0.58 (s.d. 0.33, obs. 576)	0.47 (s.d. 0.31, obs. 432)

* The first-order beliefs are elicited for all recipients.

The averages contained in Table 1 (and in the following ones) are compared by using non-parametric tests. All the statistics reported adopt the Wilcoxon signed-rank, which compares averages at the session level. Our data are independent at the session level, but not at the individual level. The Wilcoxon signed rank tests accounts for this data structure. The comparison of first-order beliefs by row (A and B respectively), taking into account low and high switching probabilities, confirms the existence of an exogenous variation in first-order beliefs. Specifically, the recipients who are matched with an out-group member under low switching probability to change partner expect that 36% of dictators will chose to roll, whereas the expected fraction increases to 45% among the recipients who are matched with an out-group member under high switching probability (i.e., 36% vs. 45%: Z=2.66, p=0.007). Similarly, the first-order beliefs of recipients who are matched under an in-group member with low switching probability to change partner assign the value 58% to the fraction of dictators who will chose to roll, whereas the recipients who are matched with an in-group member under high switching probability expect the fraction of rolling dictators to be equal to 47% (i.e., 58% vs. 47%: Z=3.06, p=0.002).

It is also worth noting that under low probability of changing partner the first-order beliefs of recipients who are matched with an in-group member are characterized by a greater average than that of recipients who are matched with an out-group member (i.e., 58% vs. 36%: Z=3.06, p=0.002), whereas under high switching probability the first-order beliefs of the two types of recipients are not statistically different from each other (i.e., 47% vs. 45%: Z=1.26, p=0.209).

Table 2. Dictators' second-order beliefs (1920 obs.) *

	Switching Probability	
	Low (12.5%) (1)	High (87.5%) (2)
Matching groups		
(A) Out Group	0.35 (s.d. 0.32, obs. 384)	0.43 (s.d. 0.32, obs. 528)
(B) In Group	0.59 (s.d. 0.34, obs. 576)	0.46 (s.d. 0.31, obs. 432)

* Second-order beliefs are elicited for all dictators before they know whether they have been re-matched or not.

Similarly, by looking in Table 2 at second-order beliefs and comparing rows A and B, under low and high switching probabilities, we can confirm the existence of an exogenous variation in expectations. Specifically, the dictators who are matched with an out-group member under low probability of changing partner believe that the recipients expect dictators to *Roll* with a probability equal to 35%, whereas the same expectation increases to 43% among the dictators who are matched with an out-group member under high switching probability (i.e., 35% vs. 43%: Z=2.74, p=0.006). The second-order beliefs of dictators who are matched with an in-group member under low switching probability to change partner are instead characterized by an expected probability to *Roll* equal to 59%, whereas this expectation is 46% among the dictators who are matched with an in-group member under high switching probability (i.e., 59% vs. 46%: Z=2.98, p=0.003).

Also note that under low probability of changing partner, the second-order beliefs of dictators who are matched with an in-group member are characterized by a greater expectation on *Roll* than that of dictators who are matched with an out-group member (i.e., 59% vs. 35%: Z=3.06, p=0.002). The same may seem to occur under high switching probability, but in this case the difference is not statistically significant (i.e., 46% vs. 43%: Z=0.78, p=0.432).

4.2 Matching, beliefs and Roll rates

Table 3 reports the averages of dictators' second-order beliefs and *Roll* rates, the standard deviations and the number of observations. The rows of the table also distinguish: (i) dictators who were matched with an out-group member (row A) from (ii) dictators who were matched with an in-group member (row B).

The table provides evidence in favor of social identity.[29] The average *Roll* rate of in-group members is higher (50%) than the average of out-group members (34%) [i.e., Z=2.98, p=0.003]. Moreover, the figures contained in the cells show that the average second-order beliefs of in-group members are higher (53%) than those of out-group members (40%) [Z=3.06, p=0.002].[30] We hence confirm the existence of a correlation between behavior and second-order beliefs (Guth *et. al.*, 2009; Ockenfeld and Werner 2014), but, in order to identify the direction of causality, we must look at exogenous variations of group membership.

Table 3. Matching, beliefs, and *Roll* rates (1920 obs.)

[29] Note that our switching mechanism tends to be unfavorable for our tests, as Ockenfels and Werner (2014) showed that there is substantially less in-group favoritism if the dictator is informed that the recipient is unaware of the shared group membership.

[30] It is worth noting that we would obtain the same results if we looked at the first-round data only (see Appendix C: Table 3C).

	DICTATORS	
	SECOND-ORDER BELIEFS	AVERAGE ROLL RATES
MATCHING GROUPS	(1)	(2)
(A) OUT GROUP	0.40 (s.d. 0.32, obs. 912)	0.34 (s.d. 0.48, obs. 912)
(B) IN GROUP	0.53 (s.d. 0.33, obs. 1008)	0.50 (s.d. 0.50, obs. 1008)

4.3 Are beliefs the determinant of in-group favoritism?

In order to understand whether *BDE* is confirmed by our experiment, the behavior of *switched* dictators must be examined when an exogenous variation of group membership is produced. To this aim we must check whether both the in-group and the out-group dictators do not change their behavior when the recipients group membership change. Table 4 contains the average *Roll* rates of switched dictators, together with the standard deviations and the number of observations.

Table 4 – Exogenous variation in group membership - Switched dictators' *Roll* rates (960 obs.).

	HIGH SWITCH PROBABILITY (840 OBS.)		LOW SWITCH PROBABILITY (120 OBS.)	
	INITIAL MATCH (RECIPIENT)		INITIAL MATCH (RECIPIENT)	
	(1) OUT GROUP	(2) IN GROUP	(3) OUT GROUP	(4) IN GROUP
FINAL MATCH				
(A) OUT GROUP	0.27 (s.d. 0.44, obs. 264)	0.27 (s.d. 0.44, obs. 192)	0.25 (s.d. 0.45, obs. 12)	0.28 (s.d. 0.45, obs. 60)
(B) IN GROUP	0.54 (s.d. 0.50, obs. 216)	0.54 (s.d. 0.50, obs. 168)	0.39 (s.d. 0.49, obs. 36)	0.75 (s.d. 0.45, obs. 12)

The eight cells/categories of the table can be considered, by row, across two dimensions: dictators who are re-matched with (i) out-group (row A) and (ii) in-group recipients (row B). These are matched by columns with other two dimensions related to the recipients' history, distinguishing those who were initially matched with (i) out-group and (ii) in-group members, under high and low switching probability (columns (1)-(2) and (3)-(4)).

It is worth noting that as recipients' beliefs only depend on their initial matching, second-order beliefs of out-group (row A) and in-group dictators (row B) are assumed to be the same in each column. According to H1, given the switching probability and the recipient's initial match (i.e., the columns of Table 4), we test whether the *Roll* rate of the switched dictators who are finally matched with out-group recipients is statistically different from that of switched dictators who are finally matched with in-group recipients. As we have two switching probabilities and two recipients' initial matches, our data allow us to carry out four tests. It must be preliminary noted that the tests based on the high switching probability

are characterized by a much higher number of observations (840 on the overall) than those based on low switching probability (120).

Results of the four tests are as follows.

1. Under a high switching probability:
   1.1. as shown in column 1, the *Roll* rates of dictators who were initially matched with an out-group member and are re-matched with an out-group member (0.27) are statistically different from those of dictators who were initially matched with an out-group member and are re-matched with an in-group recipient (0.54) (i.e., 27% vs. 54%: Z=2.98, p=0.003);
   1.2. as shown in column 2, the *Roll* rates of dictators who were initially matched with an in-group member and are re-matched with an out-group member (0.27) are statistically different from those of dictators who were initially matched with an in-group member and are re-matched with an in-group recipient (0.54) (i.e., 27% vs. 54%: Z=2.98, p=0.003).
2. Under a low switching probability:
   2.1. as shown in column 3, the *Roll* rates of dictators who were initially matched with an out-group member and are re-matched with an out-group member (0.25) are not statistically different from those of dictators who were initially matched with an out-group member and are re-matched with an in-group recipient (0.39) (i.e., 25% vs. 39%: Z=1.44, p=0.149);
   2.2. as shown in column 4, the *Roll* rates of dictators who were initially matched with an in-group member and are re-matched with an out-group member (0.28) are statistically different from those of dictators who were initially matched with an in-group member and are re-matched with an in-group recipient (0.75) (i.e., 28% vs. 75%: Z=2.62, p=0.009).

Even though Result 2.1 above seems to suggest that group identity does not always play a role, the very small number of observations which characterizes this case[31] leads us to the conclusion that, according to our test, the dictators' behavior does not generally appear to be causally affected by expectations. The test therefore provides evidence that in-group favoritism cannot be explained by beliefs and so suggests that *BDE* cannot be confirmed by our experiment (H1 is rejected). It should however be noted that beliefs might mediate the dictators' behavior within in-groups, as predicted by *BGE*.

4.4 Are in group member preferences mediated by expectations?

[31] It is the case with the smallest number of observations (12 vs. 36 distributed in 12 sessions).

We now test whether in-group favoritism is mediated by beliefs (H2), that is, whether our evidence is in favor of *GIE* or of *BGE*. In order to carry out this test, we need to use the exogenous variation in expectations, focusing on in-group dictators' behavior.

Table 5 focuses on exogenous variations of beliefs. It reports the average *Roll* rates of in-group dictators. Its cells/categories are structured along two dimensions (dictators' switching conditions and beliefs that depend on the recipients' initial matches):

a. row A focuses on non-switched dictators under high switching probability (column 2) and under low switching probability (column 4);
b. row B focuses on switched dictators who were initially matched with an outgroup or an in-group recipient under both high switching probability (columns 1 and 2) and low switching probability (columns 3 and 4).

c.

Table 5 – Exogenous variation in beliefs - In-group dictators' *Roll* rates (1008 obs.).

	HIGH SWITCH PROBABILITY		LOW SWITCH PROBABILITY	
	INITIAL MATCH (RECIPIENT)		INITIAL MATCH (RECIPIENT)	
	(1) OUT GROUP	(2) IN GROUP	(3) OUT GROUP	(4) IN GROUP
SWITCHING CONDITION				
(A) NON-SWITCHED		0.58 (s.d. 0.50, obs. 72)		0.61 (s.d. 0.49, obs. 504)
(B) SWITCHED	0.54 (s.d. 0.50, obs. 216)	0.54 (s.d. 0.50, obs. 168)	0.39 (s.d. 0.49, obs. 36)	0.75 (s.d. 0.45, obs. 12)

From row A of Table 5 we get that the average *Roll* rate of non-switched dictators who are matched with an in-group member under a low switching probability is not statistically different from that of dictators who are matched with an in-group member under a high switching probability (i.e., 61% vs. 58%: Z=0.31, p=0.753).[32] This shows that and in-group, dictators do not seem to change their behavior (we do not find a different *Roll* rate) when their second-order expectations change (i.e., under low and high switching probability). In other words, the non-switched dictators' behavior is not causally affected by their expectations.

A similar test can be carried out for row B. The average *Roll* rate of switched dictators who are matched with an in-group member under a low switching probability is not statistically different from that of dictators who are matched with an in-group member under a high switching probability (i.e., 54% vs. 75%: Z=1.40, p=0.162). Hence, switched dictators' behavior is not causally affected by their expectations.

[32] It is worth noting that a similar result holds for non-switched out-group dictators (not reported in Table 5). The average of those who are matched with an out-group member under a low switching probability is not statistically different from that of dictators who are matched with an in-group member under a high switching probability (i.e., 30% vs. 38%: Z=0.47, p=0.637).

Finally, as we have shown in Table 2 that beliefs are statistically different under a low switching probability, we can perform a further test under low switching probability. By comparing the average *Roll* rates of in-group dictators who are re-matched either with initially in-group (0.75) or initially out-group (0.39) recipients we do not find any statistically significant difference (i.e., 75% vs. 39%: Z=1.90, p=0.058). These results confirm that the switched dictators' behavior is also not causally affected by their expectations.[33]

Our conclusion is that the results from Table 5 seem to suggest that in group membership matters *per se*. We do not find evidence in favor of *BGE*, as all above tests reject H2, i.e., that group identity is mediated by beliefs. These results do not confirm those obtained by Ockenfels e Werner (2014), who find evidence supporting the view that in-group favoritism is partially belief-dependent. These findings are based on a variation in expectations which is not exogenous but based on strategic choices made by dictators in different contexts. Once an exogenous variation in expectations is obtained by using a random device, no evidence that roll rates are driven by beliefs is instead found. So, in our view, Ockenfels e Werner's (2014) results are likely to be driven more by a self-selection problem related to the strategic use of information disclosure than by beliefs.

5. Conclusions

In this paper we tried to put to a test three possible explanations of in-group favoritism: *Beliefs Driven Explanation* (*BDE*), *Group Identity Explanation* (*GIE*) and *Belief-mediated Group Identity Explanation* (*BGE*). All explanations differently combine second-order beliefs and preferences, *per se*, for people belonging to their own group, and imply different causality. In order to disentangle the effects produced by beliefs from those produced by group identity, we modified the well-known Dictator's Game by producing exogenous variations both in group membership and in beliefs. This allowed us to test whether: (i) beliefs can explain individual behavior independently of group membership; (ii) if this were not the case, individuals suffer from in-group bias independently of their expectations, so that group membership is sufficient to explain in-group favoritism, or group identity is mediated by beliefs.

[33] It should also be noted that the roll rates of non-switched dictators are not statistically different from those of switched dictators who are finally matched with recipients who were initially matched with an out-group member and were then re-matched with an in-group member (i.e., a dictator of the same color). Under a high switching probability, the test is: 58% vs. 54%: Z=0.71, p=0.480, whereas under a low switching the test is: 61% vs. 39%: Z=1.88, p=0.060. Furthermore, the roll rates of switched in-group dictators who were initially matched with an out-group member under high switching probability are not statistically different from those of switched in-group dictators who were initially matched with an out-group member in low switching probability (54% vs. 39%: Z=1.30, p=0.195).

The experiment we carried out provided the following results. First, beliefs *per se* do not seem to provide a significant explanation of in-group favoritism, so the *BDE* cannot be fully supported. Second, group identity provides a significant explanation of the phenomenon under examination. Third, in-group favoritism due to group identity does not seem to be strongly mediated by beliefs. Our experiment thus suggests to single out the *SNE* as the most powerful explanation of social identity.

These conclusions raise however further research questions, as we are now interested in understanding whether beliefs may re-enter the scene when dynamics is considered, for example when repeated interactions among members within groups,[34] or loyalty to the group is are taken into account. This represents for us a fruitful direction for future research.

Appendix A – Social identity and beliefs

In September 2017, we obtained in a five-session pilot experiment a significant correlation between behavior and beliefs conditional to the match. As shown in Table A, the second-order beliefs of the dictators who are matched with an out-group member are lower than those of the dictators who are matched with an in-group member (38% vs. 51%: Z=2.02, p=0.043) and the roll rates of the dictators who are matched with an out-group member are lower than those of the dictators who are matched with an in-group member (30% vs. 51%: Z=2.02, p=0.043). As mentioned in the main text, these correlations do not of course not imply causation.

Table A – Matching, second-order beliefs, and roll rates

	DICTATORS	
	SECOND-ORDER BELIEFS	AVERAGE ROLL RATES
MATCHING	(1)	(2)
(A) OUT GROUP	0.38 (s.d. 0.32, obs. 356)	0.30 (s.d. 0.46, obs. 356)
(B) IN GROUP	0.51 (s.d. 0.33, obs. 412)	0.51 (s.d. 0.51, obs. 412)

Appendix B – Elicitation of beliefs

The beliefs elicitation strategy is based on Vanberg (2008) and Di Bartolomeo *et al*. (2018).

Elicitation of first-order beliefs: After matching groups, recipients were informed of the value of the switching probability in their treatment (they were aware that their paired subject could be switched according to that probability) and were asked to guess what their (unknown) dictators would choose to do

[34] On the role played by repeated interactions in this context, see, e.g., Balliet *et al.* (2014) and Dorrough *et al*. (2015).

at the end of the round and thus what their payoffs would be. Recipients could make their guess by ticking one of the five-point scale described in Table B. Beliefs were then re-scaled to 1, 0.75, 0.5, 0.25, and 0. The figures reported in the main text hence represent the averages of the recipients' re-scaled responses. The payoffs correspond to a quadratic scoring rule for probability values 85%, 68%, 50%, 32%, and 15% (under the assumption of risk neutrality, quadratic scoring yields flat payoffs as probabilities approach one).

Table B – Incentives for first-order belief elicitation

The dictator will	choose *Roll*			choose *Don't Roll*	
	Certainly	Probably	Unsure	Probably	Certainly
Please tick your guess	◯	◯	◯	◯	◯
Your earnings if the dictator					
chooses *Roll*	0.65 tokens	0.60 tokens	0.50 tokens	0.35 tokens	0.15 tokens
chooses *Don't Roll*	0.15 tokens	0.35 tokens	0.50 tokens	0.60 tokens	0.65 tokens

Elicitation of second-order beliefs: Before dictators were told whether their paired subject had been switched or not, they were asked to guess his guess. Specifically, they had to guess which of the five points of Table A had been ticked by their counterpart. Correct guesses were paid 0.50 tokens.

Appendix C – First Round Results

Table C below contains the averages of dictators' second-order beliefs and *Roll* rates, the standard deviations and the number of observations reported in Table 3 of the main text but considering only the first rounds. The rows of the table also distinguish: (i) dictators who were matched with an out-group member (row A) from (ii) dictators who were matched with an in-group member (row B).

Table C. Matching, beliefs and *Roll* rates - first round only (192 obs.)

	DICTATORS	
	SECOND-ORDER BELIEFS	AVERAGE ROLL RATES
MATCHING GROUPS	(1)	(2)
(A) OUT GROUP	0.43 (s.d. 0.29, obs. 96)	0.32 (s.d. 0.47, obs. 96)
(B) IN GROUP	0.50 (s.d. 0.33, obs. 96)	0.54 (s.d. 0.50, obs. 96)

Using the rank sum test, the *Roll* rate of in-group members (54%) is higher than that of out-group members (32%) [i.e., z=3.05, p=0.002]. The same result holds by using a one-tail difference mean test based on a t-test assuming dependent samples [i.e., t=3.12, p=0.001]. The average second-order beliefs of in-group members (50%) turn out to be higher than those of out-group members (43%) when a one-tail difference mean test is employed [i.e., t=1.69, p=0.047], but not when the rank sum test is used [i.e., z=1.60, p=0.111].[35]

Appendix D – Panel regression

We use individual-level panel data to further test the different explanations of group favoritism addressed in this paper. In Table D, we estimate a random intercept logit model using GLLAMM (Stata). The logit panel regression is based on 1920 observations. The dependent variable is always the probability that the dictator chooses *Roll*. Standard errors are in brackets. By one asterisk, we indicate significance at 1% level. The other variables are not statistically significant.

The estimation considers the gender of the dictator, her/his switch condition (which is the same as that of the recipient associated with her/him[36]), and the switching probability. It also accounts for the matching: "In Group" means that the dictators and the recipients belong to the same group. "Recipient initial in group" is equal to one if the recipient was matched with a partner of the same group before the switch occurred. According to our design, "Low switching probability × Recipient's initial in group" then captures high second-order beliefs. Finally, "Round" accounts for the experiment dynamics.

Table D – Estimate of panel regressions

Male	-0.012
	(0.19)
Switch	-0.299
	(0.19)
In Group	1.544*
	(0.31)
In Group × Switch	0.141
	(0.34)
Low switching probability × Recipient's initial in group	-0.067
	(0.34)
Low switching probability × Recipient's initial in group × In Group	0.374
	(0.44)
Round	0.003

[35] It is worth noting that the available observations are insufficient to reproduce the results summarized in tables 1 and 2 of the main text – and hence in tables 4 and 5 – by considering the first rounds only.

[36] By recipient we always mean the participant matched with the dictator after the switch eventually occurs.

	(0.02)
Constant	1.334*
	(0.24)
Log likelihood	-1213.89

Estimation results are in line with the conclusions derived in the main text. Specifically, dictators who are matched with recipients belonging to the same group were significantly more likely to *Roll* and this effect was observed independently of whether their partners were switched or not. Switching did not have a significant impact on behavior both *per se* and on the dictators matched with recipients belonging to the same group. This is not surprising, as our results suggest that the group identity explanation matters *per se*.

Once we control for low switching probability and recipient's initial group we find that expectations do not matter, remember that low switching probability and recipient's initial group are associated to high expectations. This suggest that, as argued in the main text, social identity is not mediated by expectations. Expectations matter neither in general ("Low switching probability × Recipient's initial in group" is not significant) nor within the dictators who are matched with recipients belonging to the same group ("Low switching probability × Recipient's initial in group × In Group" is not significant).

References

Akerlof, G. A., and R. E. Kranton (2000), "Economics and identity," *The Quarterly Journal of Economics* 115: 715–753.

Babcock, P., Bedard, K., Charness, G., Hartman, J. and H. Royer (2015), "Letting down the team? Social effects of team incentives," *Journal of the European Economic Association*, 13: 841-870.

Bacharach, M., G. Guerra, and D. J. Zizzo (2007), "The self-fulfilling property of trust: An experimental study," *Theory and Decision*, 63: 349-388.

Battigalli, P. and M. Dufwenberg (2007), "Guilt in games," *American Economic Review*, 7: 170-176.

Battigalli, P., M. Dufwenberg (2009), "Dynamic psychological games," *Journal of Economic Theory*, 141: 1–35.

Balliet, D., Y. Wu, and C. K. W. De Dreu (2014), "In-group favoritism and cooperation: A meta-analysis," *Psychological Bulletin*, 140: 1556–81.

Bellemare C., A. Sebald, and M. Strobel (2011), "Measuring the willingness to pay to avoid guilt: estimation using equilibrium and stated belief models," *Journal of Applied Econometrics*, 26(3): 437–453.

Ben-Ner, A., B. P. McCall, M. Stephane, and H. Wang (2009), "Identity and in-group/out-group differentiation in work and giving behaviors: Experimental evidence," *Journal of Economic Behavior & Organization*, 72: 153–170.

Bracht, J., and T. Regner (2013), "Moral emotions and partnership," *Journal of Economic Psychology*, 39: 313–326.

Braver, S. (1995), "Social contracts and the provision of public goods," in *Social dilemmas: Perspectives on individuals and groups*, ed. by D. Schroeder. New York: Praeger.

Brown, J. D. (1986), "Evaluations of self and others: Self-enhancement biases in social judgments," *Social Cognition*, 4, 353-376.

Charness, G., and U. Gneezy (2009) "Incentives to Exercise." *Econometrica*, 77: 909–31.

Charness, G., and M. Dufwenberg (2006), "Promises and partnership," *Econometrica*, 74: 1579-1601.

Charness, G., L. Rigotti, and A. Rustichini (2007), "Individual behavior and group membership," *American Economic Review* 97: 1340–1352.

Chen, R., and Y. Chen (2011), "The potential of social identity for equilibrium selection", *American Economic Review* 101: 2562–2589.

Chen, Y, and S. X. Li (2009), "Group Identity and Social Preferences." *American Economic Review*, 99: 431–57.

Di Bartolomeo, G., M. Dufwenberg, S. Papa, F. Passarelli (2018), "Promise, expectations and causation," mimeo.

Dorrough, A. R., Glöckner, A., Hellmann, D.M. and I. Ebert (2015), "The development of in-group favoritism in repeated social dilemmas," *Frontiers in Psychology*, 6: 1–12.

Dufwenberg, M., and U. Gneezy (2000), "Measuring beliefs in an experimental lost wallet game," *Games & Economic Behavior*, 30: 163-182.

Ederer, F., and A. Stremitzer (2017), "Promises and expectations," *Games and Economic Behavior*, 106: 161–178.

Ellingsen, T., and M. Johannesson (2004), "Promises, threats and fairness," *The Economic Journal*, 114: 397–420.

Ellingsen, T., M. Johannesson, S. Tjøtta, and G. Torsvik (2010), "Testing guilt aversion," *Games and Economic Behavior*, 68: 95-107.

Everett, J. A. C., N. S. Faber, M. Crockett (2015), "Preferences and beliefs in in-group favoritism," *Frontiers in behavioral neuroscience*, 9.

Fershtman, C., U. Gneezy, and F. Verboven (2005), "Discrimination and nepotism: The efficiency of the anonymity rule", *Journal of Legal Studies*, 34, 371–396.

Fischbacher, U. (2007), "z–Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10: 171-178.

Goette, L., D. Huffman, S. Meier, and M. Sutter (2012), "Competition between organizational groups: Its impact on altruistic and antisocial motives", *Management Science*, 58: 948–960.

Guala, F., L. Mittone, and M. Ploner (2013), "Group membership, team preferences, and expectations," *Journal of Economic Behavior & Organization* 86: 183–190.

Guerra, G., and D. J. Zizzo (2004), "Trust responsiveness and beliefs," *Journal of Economic Behavior and Organization*, 55: 25-30.

Güth, W., M. Ploner, and T. Regner, (2009), "Determinants of in-group bias: Is group affiliation mediated by guilt-aversion?" *Journal of Economic Psychology*, 30: 814–827.

Hargreaves Heap, S. P., and D. J. Zizzo (2009). "The value of groups," *American Economic Review*, 99: 295–323.

Huang, P. H., and H. M. Wu (1994), "More order without more law: A theory of social norms and organizational cultures," *Journal of Law, Economics, and Organization*, 10: 390-406.

Ioannou, C. A., S. Qi, and A. Rustichini (2015), "Group payoffs as public signals," *Journal of Economic Psychology*, 48: 89–105.

Khalmetski, K., A. Ockenfels, and P. Werner (2015), "Surprising gifts: Theory and laboratory evidence," *Journal of Economic Theory*, 159: 163–208.

Klor, E. F., and M. Shayo (2010), "Social identity and preferences over redistribution," *Journal of Public Economics*, 94: 269–278

Liebe, U., and A. Tutic (2010), "Status groups and altruistic behaviour in dictator games," *Rationality and Society*, 22: 353–380.

Ockenfeld, A., and P. Werner (2014), "Beliefs and in-group favoritism," *Journal of Economic Behavior and Organization*, 108: 453–462.

Ostrom, E., J. Walker, and R. Gardner (1992), "Covenants with and without a sword: Self–governance is possible," *The American Political Science Review*, 86: 404-417.

Regner, T., and N. S. Harth (2014), "Testing belief-dependent models," Working Paper, Max Planck Institute of Economics, Jena.

Reuben, E., P. Sapienza, and L. Zingales (2009), "Is mistrust self-fulfilling?" *Economic Letters*, 104: 89–91.

Ross, L., D. Greene, and P. House (1977), "The "false consensus effect": An egocentric bias in social perception and attribution processes," *Journal of Experimental Social Psychology*, 13: 279–301.

Tajfel, H., M. G. Billig, R. P. Bundy, and C. Flament (1971), "Social categorization and intergroup behavior," *European Journal of Social Psychology*, 1: 149–178.

Tajfel, H. (2010), *Social identity and intergroup relations*. New York, USA, Cambridge University Press.

Tajfel, H., and J. C. Turner, (1979), "An integrative theory of intergroup conflict," in W. G. Austin and S. Worchel (eds.), *The Social Psychology of Intergroup Relations*. Brooks/Cole, Monterey: 33–47.

Tajfel, H., and J. C. Turner (1986), "The social identity theory of intergroup behavior", In: Worchel, S., Austin, W. (Eds.), *Psychology of Intergroup Relations*, Nelson Hall, Chicago, pp. 7–24.

Vanberg, C. (2008), "Why do people keep their promises? An experimental test of two explanations," *Econometrica*, 76: 1467-1480.

Whitt, S., and R. K. Wilson (2007), "The dictator game, fairness and ethnicity in postwar Bosnia," *American Journal of Political Science*, 51: 655–668.

Wetherell, M. (1996), *Identities, Groups and Social Issues*, Sage Publications, London.

Pubblicato in proprio
Dipartimento di Economia e Diritto
Facoltà di Economia
Università degli Studi di Roma “La Sapienza”
Via del Castro Laurenziano 9 – 00161 Roma

ISSN 1974-2940

Dipartimento di Economia e Diritto
Università degli studi di Roma "La Sapienza"
Via del Castro Laurenziano 9 – 00161 Roma

The Working Papers Series of the Department of Economics and Law is an electronic publication intended to allow scholars to present findings of their research activity. Submitted papers contain original, unpublished work, and represent a contribution to economic literature, from any field in economics, in the theoretical, historical and empirical perspectives. Submitted papers are subjected to double blind peer review. The Working Papers Series is catalogued in the Research Papers in Economics (RePEc) Archives, and available under Creative Commons license 3.0, Attribution- Noncommercial-No derivative work. References to the Department's Working Papers must include: the author's name, the title of the working paper and the link to the working paper.

I Working Paper del Dipartimento di Economia e Diritto ottemperano agli obblighi previsti dall'art. 1 del D.L.: 31.8.45 n. 660 e dal Decreto del Presidente della Repubblica 3 maggio 2006 n.252, art.37.