Department of Economics
Working Paper No. 265

Cognitive Ability and In-group Bias: An Experimental Study

Fabian Paetzel
Rupert Sausgruber

August 2018

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

Cognitive Ability and In-group Bias: An Experimental Study

Fabian Paetzel[a,*], Rupert Sausgruber[b]

[a] *Corresponding author: Department of Economics & FOR 2104, Helmut-Schmidt-University, Germany.*
fpaetzel@hsu-hh.de
[b] *Vienna University of Economics and Business, Department of Economics, Austria.*

Abstract

We study the role of performance differences in a task requiring cognitive effort on in-group bias. We show that the in-group bias is strong in groups consisting of high-performing members, and it is weak in low-performing groups. This holds although high-performing subjects exhibit no in-group bias as members of minimal groups, whereas low-performing subjects strongly do. We also observe instances of low-performing subjects punishing the in-group favoritism of low-performing peers. The same does not occur in high-performing or minimal groups where subjects generally accept that decisions are in-group biased.

JEL: C92, D31, D63

Keywords: cognitive ability, group identity, entitlements, social preferences, minimal groups, punishment, social norms, social status

[*] We thank Meike Benker, Lars Feld, Sarah Necker, Ulrich Schmidt, Thomas Stephens, Stefan Traub, James Tremewan, the participants at the ESA international conference in Jerusalem 2016, the ESA European meeting in Bergen 2016, the GfeW conference in Giessen 2016, the micro-empirical workshop in Lüneburg 2017, the annual meeting of the research group FOR2104 in Bremen 2017 and the VfS Annual Conference 2017 in Vienna for helpful comments, discussions and encouragement.

1. Introduction

In today's societies, groups are often occupation specific. As such, they are congruent with organizational structures, to exploit the benefits from specialization and agglomeration. Groups therefore tend to differ along performance-related dimensions, such as required skills, economic returns from effort, and incentives for effort-based selection. Moreover, economic performance tends to be correlated within naturally occurring groups, for instance, because of structural features of local housing markets, determinants of migration, and the fact that learning and the spread of information are bounded by groups. Such correlations imply that individuals are more similar within than across groups, thereby contributing to the salience and comparability of performance differences between groups. Economists have long suspected that these factors shape performance perceptions and subsequent decisions.[1]

In this study, we explore the role of performance differences in a task requiring cognitive effort in a laboratory experiment on in-group bias. In-group bias is defined as choosing in favor of in-groups at the expense of out-groups (for a review, see Hewstone et al., 2002). There is a large and rapidly emerging literature showing that group membership affects individual behavior in a wide range of situations. However, existing results about in-group bias are often inconsistent across the literature. Standardized effect sizes estimated from meta-studies are generally quite small, but the numbers vary substantially conditional on the type of the study (see, Balliet et al., 2014; Lane, 2016).[2]

Our study supports this general impression and shows that in-group bias depends on both subjects' own performance and on how they perceive the performance of other members of the in-group, relative to an out-group. We show that cognitive ability affects in-group bias in performance-based groups and in minimal groups. We measure the in-group bias by means of a non-strategic allocation choice between in- and out-group members. Our most prominent result is that in-group bias is strong in groups consisting of high-performing members, and weak in low-performing groups. This result obtains even though high-performing subjects are not, on average, more or less other-regarding than low-performing ones, i.e., the asymmetry effect shows up in terms of a higher variance rather than a difference in means. As a second result, we find that high-performing subjects exhibit no in-group bias as members of minimal groups, whereas low-performing subjects strongly do. These two findings show that the asymmetric bias between high- and low performing groups dominates and overturns differences at the individual level between how subjects react to being categorized as members of random (minimal) groups. Finally, we explore whether the strength of a social norm, as measured by norm enforcement, differs across treatments. We observe instances of

[1]See, for instance, Jackson (2011) on the study of networks generally from an economists perspective, Luttmer (2005) on relative earnings between neighbors and well-being, and Alesina et al. (1999) on social fragmentation and the provision of public goods. For theoretical arguments as to why beliefs might be endogenous to empirical levels of performance, see Piketty (1998) and Alesina and Angeletos (2005).

[2]Balliet et al. (2014) and Lane (2016) report effect sizes between a quarter and a third of a standard deviation, respectively. Restricting to observations finding a significant bias, Lane (2016) reports an average effect size of two thirds of a standard deviation.

low-performing subjects punishing the in-group favoritism of low-performing peers. The same does not occur in high-performing or minimal groups where subjects generally accept that decisions are in-group biased. Together, these results suggest that perceived differences in performance give rise to an asymmetric deference relationship between groups.

In our design, we observe a simple choice between two allocations: the decision maker can implement either an equal or unequal distribution of payoffs among three subjects including herself. The unequal distribution pays more to the decision maker and another subject, however, at the expense of a third subject. To measure the extent of in-group bias, in two conditions, A and B, the design randomly varies which of the other subjects belongs to the same group as the decision maker. In condition A, the materially self-interested choice inflicts a loss on the in-group member. Condition B is identical except that the self-interested choice hurts the out-group member. This design provides a measure of in-group bias as the difference in average behavior between conditions A and B.

In the main treatments of the experiment, we assign subjects to either high- or low-performing groups based on their scores on a cognitive ability test. Group assignment is common knowledge. As a control, we implement minimal groups based on a criterion unrelated to performance. Therefore, decision makers in the experiment can be members of any three "identity groups:" high, low, or minimal. For each of these groups, we measure the in-group bias as the average treatment effect between conditions A and B, thereby accounting for individual-specific unobserved factors along with subjects' assignment to performance-based groups. Finally, we repeat the experiment among the same subjects with the only difference being that the formerly passive member of the same identity group is given the opportunity to punish the decision maker.

We regard our findings as important, as most natural groups differ with respect to income, wealth, and educational achievements. Evidence from surveys illustrates that differences in performance due to work preferences and abilities are indeed highly salient in people's perceptions (see Alesina and Giuliano, 2011). For example, Fiske et al. (2002) observe that Americans generally associate high group status with favorable traits, such as being competent or deserving, and low group status with laziness or not being intelligent. Our paper contributes to understanding possible consequences of perceived performance of groups. We thereby provide a fruitful synopsis of two prominent, but so far largely separated lines of the literature, namely, research on social identity and group status in social psychology, and the work on social preferences and entitlements in economics (see section 2 for references and a thorough discussion of the literature).

Our results confirm the social identity approach in that we, as many others, already observe a significant in-group bias in minimal groups. The contribution of our paper is to show that group performance influences the in-group bias. Two consistent explanations are that (i) subjects derive a stronger sense of membership in high-performing groups and that (ii) observable performance differences give rise to equity considerations between groups. These explanations likely hold simultaneously, and they are both valid for modeling asymmetric social norms in groups of different performance.

Our favored explanation is that social identity matters, and that entitlement motives add to the in-group bias in high-performing groups, but work in the opposite direction in low-performing groups. The underlying reason is that the members of low-performing groups view their peers as less deserving than members of high performing groups. This interpretation implies that the members of low-status groups implausibly perceive themselves as inferior, irrespective of how the difference in social status came about. Indirect support comes from evidence suggesting that a sense of "consensual inferiority," amply observed among low-status groups, breaks down if status differences are perceived as illegitimate in the sense that they arise from factors beyond an individual's control (see, for example, Ellemers et al. (1999), Bettencourt et al. (2001), and Levin et al. (2002); for a more general overview of the research on legitimacy in psychology, see Tyler (2006)). Suggestive evidence also comes from our experiment. In our design, the members of a low-performing group know that the respective out-group member has worked hard in relation to the in-group. That some of them punish their peers for being in-group biased is therefore consistent with a concern for entitlements. We note, however, that the usual understanding of social status is often intertwined with perception of entitlements.

The rest of the paper proceeds as follows. In the next section, we explain how our study relates to the existing literature. In Section 3, we explain the experimental design and how we measure in-group bias. In Section 4, we report the results, and Section 5 concludes and discusses further implications of our study.

2. Further Literature

The literature on in-group bias in economics starts with Akerlof and Kranton (2000). Following the social identity approach in social psychology (Tajfel and Turner, 1979; Turner et al., 1987), these authors define identity or self-image as derived from prescriptions of appropriate behavior for different social categories or groups. In this model, people are reluctant to deviate from group-specific norms, which implies that their behavior is biased toward the categories or groups to which they belong. Most of the empirical research in economics focuses on this immediate implication of the model and tests whether subjects exhibit different behaviors toward in- and out-group members, as we do. This literature illustrates that decisions are in-group biased in a large class of games (among others, see Hargreaves Heap and Varoufakis, 2002; Charness et al., 2007; Chen and Li, 2009). To mitigate systematic differences in unobserved characteristics across groups, most studies randomly assign subjects to artificial minimal or near-to-minimal groups. In comparison to our study, this literature does not focus on the asymmetry of in-group bias between subjects or groups with different characteristics.

Literature on heterogeneity in in-group bias. In the economics literature, only a few papers deal with the question of how in-group bias varies with the characteristics of individuals or groups. Klor and Shayo (2010) study voting on redistribution among subjects from two distinct natural groups. They find pronounced group bias on the aggregate level, but this effect is driven by only a third of subjects, who are classified as "social identifiers." Further details of their study suggest that social identifiers are different outside the lab, too. In particular, Klor and Shayo (2010) find from survey data that the correlation between income and redistributive preferences is weaker for social identifiers than for materially self-interested subjects.[3] Kranton et al. (2016) use a within-subject design to study group bias in a minimal-group and a political-group treatment. They find that about a third of subjects show no in-group bias in either treatment, and subjects who are in-group biased in the minimal-group treatment are more likely to be biased in the political-group treatment. They also observe that subjects who have no party affiliation are least likely to be in-group biased. In a companion study on M-Turk, Kranton and Sanders (2017) test for correlates of individual propensities to treat people differently. Although they find no effect of demographics and personality, their results suggest that subjects coming from regions that have experienced a drop in economic performance are the most likely to be social identifiers. These results confirm the need to account for the possibility that subjects propensities to treat people differently correlate with individual performance, as we do in our study. Moreover, recent evidence provided by Benjamin et al. (2013), Hoppe and Kusterer (2011), and Abeler and Marklein (2016) suggests that individuals with low cognitive ability are more prone to behavioral biases in individual decision making. To our knowledge, there is no study considering a correlation between cognitive ability and in-group bias in minimal groups.[4]

The closest match to our paper that we are aware of is Hett et al. (2016). In an experiment similar to Hargreaves Heap and Zizzo (2009), these authors measure the value of groups as subjects' willingness to accept (WTA) being assigned to another group, as compared to staying in their own group. While Hargreaves Heap and Zizzo (2009) use minimal groups to identify the pure psychological benefits of belonging to a group, Hett et al. (2016) assign subjects to "real" groups that differ by whether they study at the same or a different university, and by whether their performance in a summation task is high or low. Hett et al. (2016) find that the stated WTAs are highest in high-performing groups at a subjects' own university. Furthermore, they find that the WTAs are associated with in-group bias in a dictator game. Therefore, their results are similar to ours, in that they also observe a high level of in-group bias in high-performing groups. There are also important differences between Hett et al. (2016) and our study, however. They focus on measuring the preference for belonging to a group, and not so much on a general relation between in-group bias and performance, and their design does not contain a minimal-group treatment to test whether subjects of different performance differ with respect to a general bias toward an out-group per se. Rather, in their design, every group is characterized by variation along two real dimensions that are not necessarily orthogonal, and their subjects work on a joint quiz to enhance group identity. Correlates of general biasedness and idiosyncratic influences of groups may be difficult to disentangle, because of these features of the design. In contrast, we aim to identify the effects of performance differences on in-group bias in the absence of any potentially intervening forces.

[3]Strong national identification among poor voters as observed from surveys, may then explain why they exhibit lower-than-expected support for redistribution (see Shayo, 2009).

[4]There is, however, a literature focusing on individuals' performance in cognitive and social tasks as the outcome, rather than the cause, of in-group biased behaviors. See, e.g., Schmader et al. (2008), for a study on the effects of negative stereotyping.

Social psychological literature on in-group bias and status. The literature in social psychology defines social identity as a sense of self derived from membership of a group and the characteristics shared with its members (Tajfel and Turner, 1979). Social identity theory assumes that the in-group bias follows from a sequence consisting of social categorization (labelling someone as member of a group), social identification (deriving a sense of self), and group comparison (deriving a higher sense of self from favorable group comparison).[5] With regard to the latter, an extensive body of research from social psychology suggests that in-group bias is higher for high-status groups. The term group status is thereby used broadly, as perceived differences between groups on valued dimensions of comparisons.[6] A prominent strand of the literature considers group bias of ethnic groups. This literature reports evidence in support of in-group bias of high-status majorities and out-group favoritism of underprivileged minorities (for a review see Tajfel, 1982). However, and in line with our study, these effects have been observed to depend on whether individuals accept status differentials as legitimate in the sense of being equitable and just (for meta analyses see Mullen et al., 1992; Bettencourt et al., 2001). Although patterns of this kind would also have important implications in the domain of economics, the literature has hitherto not dealt much with the role of status differences in in-group bias.[7]

The social psychology research on group status is closely related to our study in that group status is often manipulated by assigning individuals to groups on the basis of performance in a task.[8] It differs from our research, however, both in terms of focus and methods. Social psychology focuses on the study of ethnocentrism understood as a positive evaluation of the in-group relative to the out-group. As a result, it is overwhelmingly focused on the study of stereotypes and prejudice, as captured by self-reported, survey-based measures of perceived traits and evaluations. For example, Mullen et al. (1992) in their meta-analysis only include studies that measure perceptions. When it comes to studying discriminatory behavior, the prevailing method consists of so-called *other-other-allocations* in which decision makers can transfer resources between in-group and out-group members, at no cost to themselves. This method was pioneered by Tajfel (1970), to illustrate that there is discrimination against an out-group "even if there is no reason for it in terms of the individual's own interests." To the best of our knowledge, there is no controlled study on

[5]Self-categorization theory extends this approach by focusing on the interplay between multiple layers of identity, in particular, personal identity as derived from interpersonal comparisons and social identity as gained from group membership (Turner et al., 1987). An often-stated assumption is that individuals with high self-esteem from interpersonal comparisons have less of a need to engage in group comparisons. Although the details of such correlations are the subject of considerable debate (see, e.g. Aberson et al., 2000), this sort of reasoning suggests that in-group bias likely also depends on individual characteristics.

[6]Next to status, social psychology considers similarity among group members as a second main driver for enhancing group identity (Tajfel and Turner, 1979). Similarity by itself is, however, not sufficient to explain an asymmetry effect.

[7]See, however, Shayo (2009) and Holm (2016) for models of jurisdictional identification depending on the status of groups. Tsutsui and Zizzo (2014) and Butler (2014) experimentally manipulate status by only changing the label of a group. Different from us, these studies consider behavior in strategic environments.

[8]The nature of the tasks used for this purpose differs widely across the literature. See Cheng et al. (2014) for a review of experimental and survey measures to manipulate status in social psychology research. Some even employ tests of cognitive ability, as we do (Brewer et al., 1993).

the relation between in-group bias and status that includes costly decisions.[9] Another prominent question is whether status associations are automatic in the sense that they underlie unconscious attitudes and beliefs (for recent examples, see Rudman et al., 2002; Newheiser and Olson, 2012).

Compared to this literature, we follow the economic approach and measure bias as revealed preference rather than a difference in perceptions. This is of considerable value because individuals may not perceive groups differently, may be wrong in their perceptions, and because perceptions may have no impact on behavior (for a general discussion of these and other concerns with attitudinal measures, see Bertrand and Mullainathan, 2001). Moreover, in the absence of real incentives, the economic prediction is that subjects are indifferent, which gives rise to potential concerns that the results could be artifacts of context and experimenter demand. In line with this, Chen and Li (2009), as part of a comprehensive study on in-group bias, evaluate common methods to induce group categorization in social psychology research. They find that some of the methods commonly used to induce group categorization significantly increase attachment to groups as measured from a questionnaire, but have no effect on behavior.[10]

There are also conceptual differences between evaluations and behavior as measures for in-group bias. To see this, assume that people agree about the status distinction between groups such that evaluations would not be biased.[11] For example, people might agree that high- and low status groups are not equally deserving. Such beliefs might matter for behavior nonetheless (see the literature on entitlements below).

Another fundamental difference is the use of deception. Experimenters in social psychology often use bogus procedures, such as false feedback about test performance, to manipulate group status. While this has the advantage of keeping group assignment essentially random, it comes at the danger of losing experimental control (Ortmann and Hertwig, 2002). We do not use deception in our design.

Literature on entitlements. Social identity theory is not the only explanation for higher bias in high-performing groups. An alternative, which has been largely ignored in the literature on in-group bias, is that the asymmetry effect arises mechanically, from a meritocratic notion of entitlement in situations in which performance is clustered in groups. An extensive body of research finds that social perceptions regarding the fairness of relative positions in income and wealth depend on the

[9] In a follow-up to Mullen et al. (1992), Bettencourt et al. (2001) extend the inclusion criteria for the meta-analysis to reward allocations, resource allocations, and Tajfels matrices as the dependent variable. None of these are costly to the decision maker. In a recent study, Hays and Blader (2017) consider the effect of status on dictator giving. In close proximity to the research in economics on entitlements (see below), high-status dictators give less to low-status receivers only when status is earned by performance in a task. This study is not about in-group bias, but it confirms our impression that, in social psychology the effect of status on pro-social behavior is generally largely unexplored.

[10] Partly in line with this interpretation, research in social psychology finds only a weak relationship between explicit and implicit (unconscious) measures of group bias (for a discussion see Hewstone et al., 2002).

[11] For an illustration, see Brauer (2001). This author proposes an innovative design in which subjects evaluate several out-groups on a range of attribute dimensions. Although the results confirm strong stereotypes among occupational groups, they provide no proof of a stronger evaluative bias for high-status groups. The reason is that the perceived differences are largely consensual between high- and low-status subjects. These findings contrast with social identity theory according to which members of low-status (high-status) groups try to positively differentiate the own group on dimensions unrelated (related) to status (see Mullen et al., 1992; Bettencourt et al., 2001).

extent to which individuals are perceived as accountable for differences in economic performance (see, Konow, 2000; Fong, 2001; Croson and Konow, 2009; Gill and Stone, 2010; Krawczyk, 2010; Cappelen et al., 2013; Gill and Stone, 2015; Mollerstrom et al., 2015).[12] If individuals acknowledge entitlements and, at the same time, performance differs across groups, these two assumptions have the straightforward, but nonetheless surprising, implication that high-performing individuals in high-performing groups have a greater tendency to favor the in-group over the out-group.

Further evidence in support of the observed asymmetry in terms of entitlements comes from surveys and experiments pointing to a relationship between performance-related traits and what is typically termed as a person's (rather than a group's) status or prestige (see Weiss and Fershtman (1998) and Heffetz and Frank (2008) for a thorough discussion of possible meanings of status in economics, including its function as a signaling device). Ball et al. (2001, p. 161) define status as a "ranking in a hierarchy that is socially recognized and typically carries with it the expectations of entitlements to certain resources." In a series of experiments including different games, Ball et al. (2001), Hoffman et al. (1994), Hoffman and Spitzer (1985), and Ball and Eckel (1998) find that subjects who earn their role in a game by the investment of effort are viewed as deserving a greater claim to economic rewards. This result indicates that individuals agree on relative positions based on productivity-related traits (for an evolutionary explanation, see Henrich and Gil-White, 2001). Such agreement legitimizes an asymmetric relationship according to which low-performing individuals pay deference to high-performing ones, and a person of high performance expects a reward. Perhaps surprisingly, this literature has not dealt with implications of status for in- and out-group behavior.

Further related research. Several additional papers relate to our research. In terms of methods, our study relates to a recent literature on the role of cognitive ability in economic decision-making (see, e.g., Burks et al., 2009; Rustichini, 2015). Similar to us, Gill and Prowse (2016) and Proto et al. (2016) classify subjects as either high- or low-cognitive ability in experiments. Their studies investigate the role of cognitive skills in repeated strategic interactions. In a similar vein, Hanaki et al. (2015) show that diversity in cognitive ability increases mispricing in markets. These studies, however, are not about group bias. Generally, cognitive ability is a major determinant of important life outcomes (see, e.g. Heckman et al., 2006). The literature in economics has only started to explore the mechanisms that may underlie this relation. Next to a direct effect that already matters for individual, non-strategic economic decisions, differences in cognitive ability play a role through their effect on the level of strategic sophistication and beliefs.

On a more general note, our study is related to the literature on group identity and social preferences (see Chen and Li, 2009; Klor and Shayo, 2010; Lindqvist and Östling, 2013). To our knowledge, our paper is the first to discuss the implications of entitlement considerations for in- and out-group behavior within that field. In addition, our study is related to the emerging literature

[12]For an overview of different fairness principles, compare e.g. Miller (1999), Konow (2003), Cappelen et al. (2007), and Nicklisch and Paetzel (2018). Compare Homans (1974) and Selten (1988) for earlier contributions on equity theory.

on in-group bias in "real" social groups as opposed to artificially assembled "minimal" groups (see, Hewstone et al., 2002; Bernhard et al., 2006; Fowler and Kam, 2007; Goette et al., 2012; Cappelen et al., 2013; Schniter and Shields, 2014; Chowdhury et al., 2016). Instead of sorting subjects into new groups, several recent papers in economics use priming techniques to study social identity in natural groups (among others, see Benjamin et al., 2010; Mobius et al., 2016). These studies typically evaluate the priming intervention under stable conditions against a neutral control. Similar to comparing two conditions A and B in an experiment, as we do, this method circumvents many difficulties and identifies the causal impact of identity in heterogeneous social groups (for a discussion, see Cohn and Maréchal, 2016). The results of this literature suggest that there are important quantitative moderators of in-group bias. Our study complements this literature by hinting at performance perceptions as an important source of heterogeneity in in-group bias between naturally occurring real groups.

Finally, we add to a recent literature studying how hierarchical systems are sustained by social norms and their enforcement (see, Hoff et al., 2011; von Essen and Ranehill, 2013; Falk, 2017). Among others, a difference between these studies and ours is that they consider different dimensions of status from those that we do.

3. Experiment

The experiment has four stages: a performance stage, a group-assignment stage, and two decision stages (see Figure 1). The subjects receive instructions separately for each stage (compare the instructions in the Appendix). At any particular stage, the subjects are not yet informed about what will happen in the subsequent stage(s).

The performance stage. In this stage, we ask the subjects to answer a series of questions, which have the format of nonverbal multiple choice questions commonly used in tests of cognitive ability.[13] Subjects receive no payment for performance. Therefore, next to cognitive ability the test scores are likely to reflect characteristics associated with effort (Segal, 2012). At the end of this stage, all subjects receive private feedback on their scores.

Depending on the group assignment (described in detail below in this section), some subjects additionally learn whether their own score falls in either the upper (high) or lower (low) half of the distribution of scores in the same session.

The group-assignment stage. We use two methods of assigning subjects to social groups. We have treatments in which we assign subjects to "performance-based groups," based on their scores from the performance stage. In particular, we split them by session medians into two groups: the group that scores "High" and the group that scores "Low" in every session. Subjects in these treatments are labeled as being of type "Low" or "High."

[13] We measure cognitive ability in a test used by Putterman et al. (2011) and Kamei et al. (2015). The test uses questions that are taken from de Séréville and Myers (1994), and are based on Raven's progressive matrices. Each question offers eight possible answers, only one of which is correct. Subjects are given one minute per question. An example is provided in the instructions. See the Appendix: Figure B.1.

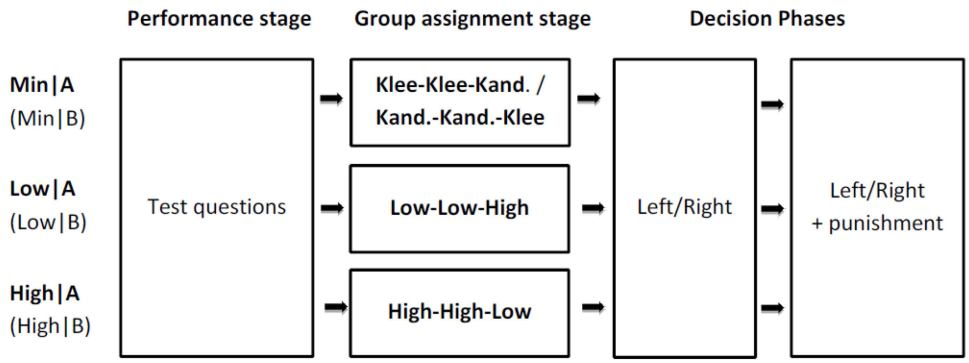


Figure 1: Overview and timeline of the experimental design

As a control, we run treatments in which we assign subjects to "minimal groups." We follow the method proposed by Chen and Li (2009). Accordingly, subjects view five pairs of paintings. In each pair, one painting is by Wassily Kandinsky and the other by Paul Klee. Subjects indicate which painting they prefer in a given pair and subsequently, are split by the median preference of subjects in the same session. This method divides the subjects into two groups, the "Klee" group and the "Kandinsky" group.[14]

Treatments and the first decision stage. After all subjects were assigned to either the performance-based or the minimal groups, we further match them into subgroups of three. The matching is such that, in all subgroups, subjects take one of three roles: (i) an active in-group member who is the only subject that makes a decision (I_A), (ii) a passive in-group member who does not make a decision (I_P), and (iii) a passive out-group member (O). The group assignment and the manner of matching subjects into subgroups generates four decision treatments.

Figure 2 illustrates the treatments. Consider first the minimal-groups treatments in the upper part of the Figure. In the treatments labeled Klee|A and Klee|B, there are Klee–Klee–Kandinsky groups in which the decision makers are of type Klee; likewise, in Kand.|A and Kand.|B, there are Kandinsky–Kandinsky–Klee groups in which the decision makers are of type Kandinsky (the details on conditions A and B are explained later in this section). Due to the minimal group paradigm, decisions should not differ between Klee|A and Kand.|A or between Klee|B and Kand.|B. Hence, we regard the decisions from these groups as taken from the same treatments, labeled as Min|A and Min|B.[15]

Next, consider the performance-based groups in the lower part of the Figure 2. In these treat-

[14]Subjects' performance in the cognitive-ability test does not differ between performance-based treatments and minimal treatments ($\chi^2 = 15.3961$ with $p = 0.352$). Per design, the test scores differ significantly between subjects of type LOW and HIGH ($\chi^2 = 92.6$ with $p < 0.001$).

[15]Our results show there is indeed no difference between the Klee and Kandinsky types of decision makers. See section 4.

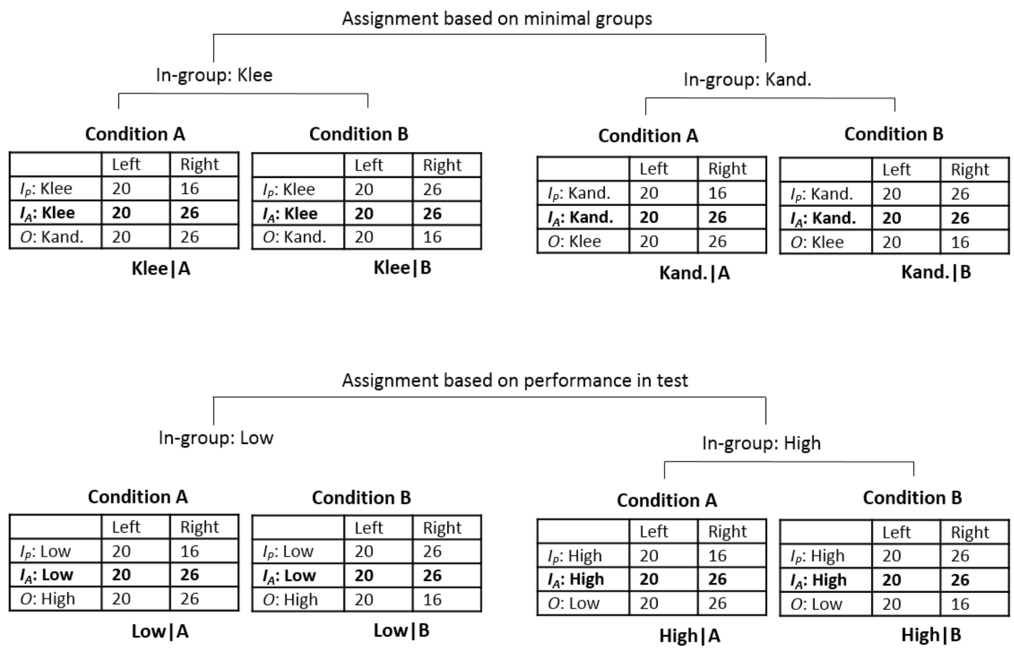


Figure 2: Overview of treatments, roles, and distributions.

ments, 1/3 of all High-scoring and 1/6 of all Low-scoring subjects of a session are randomly assigned to High–High–Low groups with a High-scoring subject in the role of the decision maker; the remaining subjects (1/6 of all High-scoring and 1/3 of all Low-scoring subjects) are matched to Low–Low–High groups with a Low-scoring decision maker. In this way, we generate one treatment in which the decision makers score high in High–High–Low groups (see High|A and High|B in the lower panels of Figure 2), and another treatment in which they score low in Low-Low-High groups (Low|A and Low|B).

The decision. There is just one decision maker (I_A) in every group. The task of the decision maker is to choose between two earnings distributions, LEFT and RIGHT (compare Figure 2). The LEFT choice implies that all group members receive the same payoff. In condition A, if the active in-group decision maker (I_A) chooses RIGHT, the incomes of I_A and the out-group member (O) increase, and the income of the passive in-group member (I_P) decreases. Therefore, by opting for LEFT the decision maker makes a costly decision in favor of the in-group member, at the expense of the out-group member. Figure 2 illustrates this choice for the parameters we use for the experiment (see panels labeled "Condition A"): if I_A chooses LEFT, all three group members earn 20 points; if this subject chooses RIGHT, I_A and O earn 26 points, whereas I_P earns 16 points.[16] The decision

[16] As an example, consider the choice of an employer between two job candidates who differ along two dimensions: productivity and identity. For employers at the margin of indifference, choosing the in-group candidate would be costly in terms of productivity differences and would actively discriminate against the out-group candidate. Lower productivity is considered by making the distribution LEFT less efficient with respect to the sum of payoffs (following the Kaldor–Hicks criterion).

maker only makes one choice. After that, the first decision stage is over.

A measure of in-group bias — "Condition A*" versus "Condition* B*".* In "Condition *A*," next to being biased toward the member of the in-group, subjects might choose LEFT because they prefer an egalitarian payoff distribution irrespective of any in- and out-group considerations. Moreover, in performance-based groups, decision makers in High|A and Low|A might differ regarding social preferences and other unobserved traits, because of selection. To net out these effects, we implement treatments labeled as "Condition *B*," which swaps the payoffs between I_P and O (see the respective panels in Figure 2). For example, between High|A and High|B, subjects in the role of I_A face identical decisions in terms of their own payoffs and the overall distribution of payoffs among themselves and others. Therefore, if their decisions were independent of the in- and out-group dimension, we would observe no systematic difference between these two conditions. By contrast, if decisions were biased in favor of the in-group member, we would observe more Left choices in High|A than in High|B. The same argument applies for the differences between Low|A versus Low|B, and Min|A versus Min|B. Therefore, the difference-in-difference effects provide a measure of in-group bias for different identity groups.[17]

Decision mode. Subjects are randomly assigned to their roles, given the constraints of the group assignment stage. The roles are fixed for the entire experiment. However, to increase the number of observations per group, at the time of making the decision, the two in-group members I_A and I_P do not yet know which role they are in. These subjects make the decision conditional on being in the role of I_A, and they only learn their actual role at the end of the experiment.[18]

In-group punishment and the second decision stage. If subjects were concerned about a norm of loyalty within the group, they would be willing to sacrifice money to enforce it (see Harris et al. (2015) for evidence of group favoritism as a social norm). Moreover, our interest is to observe whether the strength of a social norm, as measured by norm enforcement, differs across treatments. To allow for this option, we introduce a second decision stage. This stage is identical to the first one, with the exception that subjects in the role of I_P can now assign up to 4 deduction points to subjects in the role of I_A. For subjects in the role of I_P, sending 1 deduction point costs 1 point in their own earnings. For those in the role of I_A, each punishment point received reduces earnings by 3 points.

In the second decision stage of the experiment, the two in-group members first enter their decision conditional on being in the role of I_A and then, before they learn their true roles, they choose their deduction points conditional on being in the role of I_P and on whether I_A has chosen LEFT or RIGHT. At the time of deciding between LEFT and RIGHT, subjects know that their

[17] A different approach to account for selection is used by Ball et al. (2001). In their design, subjects performed a quiz and were sorted into high-status ("Star") or low-status ("No Star") groups. These groups were actually formed independently of performance, but the instructions were written in a way that made subjects believe that the members of the Star group were deserving.

[18] Note that while asking subjects to put themselves in each other's roles might in itself generate a feeling of group attachment, we apply the same procedure in all treatments. Therefore, such effects would difference out across treatments. See also Brandts and Charness (2011) for a discussion of using the "strategy method" in experimental economics.

choice might be punished by their in-group peer.[19] After all subjects have completed their choices, they learn their roles and the respective decisions are implemented.

We only hand out the instructions for the second decision stage at the end of the first decision stage. At this point, the subjects do not yet have any information about the decisions and outcomes of the first decision stage. It is still possible that decisions change following some systematic pattern over repeated decisions. However, we are interested in the norm-enforcement behavior of subjects in the role of I_P conditionally on the behavior of subjects in the role of I_A. Order effects between phases are of no relevance to that question.

Discussion of the design. Our interest is to test whether the magnitude of the in-group bias depends on individual characteristics as well as on how subjects are assigned to groups. In the treatments with minimal groups, group assignment is orthogonal to performance. This treatment enables us to explore how subjects of different cognitive ability react to being categorized in minimal groups. Other than that, we expect the in-group bias to be symmetric between the Klee and Kandinsky types.

For the reasons given in sections 1 and 2, we expect the in-group bias to differ between high- and low-performing groups. In High|A, subjects in the role of I_A know that they are of type "High." Hence, subjects who are members of high-performing groups may derive a stronger sense of self from the favorable group comparison. Moreover, they might consider others, who are of type "Low," as less deserving. In this case, entitlement considerations would mitigate the difference between High|A and High|B. The opposite effect should be obtained for decision makers of type "low" between conditions Low|A and Low|B.

As noted above, our design provides a difference-in-difference measure of in-group bias for different performance groups. This method works around issues of selection bias and unknown third variables that may characterize the decision makers in performance-based groups. A remaining concern is that the very belief by decision makers in the existence of any criterion that correlates with performance may affect the results. Consider beliefs about income and gender as two plausible examples: subjects could believe that individuals who score high in a test of cognitive ability have higher levels of income and wealth. Similarly, they could hold stereotypical beliefs according to which men dominate in high-performing groups.[20] Given such beliefs, subjects motivated by outcome-based social preferences supposedly would want to distribute payoffs from high- to low-performing subjects because they regard the latter as less wealthy and more in need. Moreover, existing evidence from gender discrimination experiments shows that there is significant favoritism toward the opposite gender (see, Lane, 2016). Both arguments run contrary to our expectation of

[19] In addition, we ask the out-group members O to state their expectations about how many subjects in the role of I_P would punish I_A. We ask: *"Out of 10 subjects in the role of I_P, how many do you think will send 'deduction points' to I_A for choosing LEFT?" and "Out of 10 subjects in the role of I_P, how many do you think will send 'deduction points' to I_A for choosing RIGHT?"* Subjects are paid additional 5 points if they guess either one of these answers correctly.

[20] Such beliefs indeed seem warranted, according to empirical studies: see Heckman et al. (2006) on the relationship between cognitive ability and labor market outcomes, and Croson and Gneezy (2009) for a survey and discussion of gender differences in relation to confidence in gender-neutral tasks.

stronger in-group bias in high-performing groups, which suggests that the beliefs argument would bias the results in the reverse direction.[21]

With respect to the second decision stage, our main interest is to see whether the punishment behavior reflects any asymmetric pattern of in-group bias across different identity groups. By choosing the materially selfish option (RIGHT), a decision maker inflicts a loss on the passive in-group member in condition A but not in B. Hence, we expect that (i) the in-group members punish the decision makers more strongly for choosing RIGHT in conditions A than B, and that (ii) this asymmetry varies by the extent of in-group loyalty across groups.

Payment and number of observations: At the end of the experiment, we calculate subjects' earnings from both decision stages and pay them at the exchange rate of 1 point = €0.35.

We ran the experiment at the Vienna Center for Experimental Economics in October 2013. In total, 246 subjects participated in 12 sessions. There were between 18 and 24 subjects in each session. A session lasted for approximately 1 hour, the average subject earned €15.1, and no additional show-up fee was paid. Subjects were recruited via ORSEE (Greiner, 2015) and the experiment was programmed and conducted with z-Tree (Fischbacher, 2007). Table 1 shows the number of groups and observations per treatment. Because the design elicits the decisions of both in-group members I_A and I_P in the strategy mode, there are two observations per group.

Table 1: Treatments and numbers of observations

Treatment	subjects	no. of groups (no. of observations in brackets)
Low|A	42	14 (28)
Low|B	39	13 (26)
High|A	42	14 (28)
High|B	39	13 (26)
Min|A	48	16 (32)
Min|B	36	12 (24)
Total	246	82 (164)

Notes. The number of observations is twice the number of groups because two subjects of the same type in every group make a decisions conditional on being in the role of I_A.

[21] Babcock et al. (1995), Messick and Sentis (1983), and Rodriguez-Lara and Moreno-Garrido (2012) show that subjects tend to act on different fairness principles in a self-serving manner. Note that this concern is unlikely to affect our measure of in-group bias; for example, high-performing decision makers who put more weight on a norm of equity instead of equality because of being self-serving would choose RIGHT in both conditions, A and B.

4. Results

We find clear evidence for an in-group-bias over the whole sample. On average over all treatments, 55.68% (49/88) of the subjects chose LEFT in condition A. In contrast, only 28.94% (22/76) of the subjects chose LEFT in condition B. Accordingly, the average in-group bias amounts to 26.74 percentage points. This difference is highly significant (55.68% vs. 28.94%, $p = 0.001$, χ^2 test, two sided).

Next, we firstly focus on whether subjects who score high in performance differ in terms of in-group bias from subjects who score low. After that, the analysis compares the behavior between performance-based and minimal groups for both high- and low-performing subjects. Finally, we analyze the punishment behavior in the second decision phase of the experiment.

4.1. Ingroup-bias and group performance

Figure 3 shows the relative frequency of LEFT choices for the performance-based treatments in the first decision phase. First, consider the behavior of low-scoring subjects in the two leftmost bars: 57.14% (16/28) of low-scoring subjects in the role of I_A chose LEFT in treatment Low|A. While this number indicates a great deal of solidarity with the losing subject of the group, only a small part of the result can be attributed to in-group bias. In Low|B, 42.31% (11/26) of low-scoring subjects chose LEFT. Hence, the in-group bias amounts to 14.83 percentage points. This number is not significantly different from zero (57.14% vs. 42.31%, $p = 0.276$, χ^2 test, two sided).

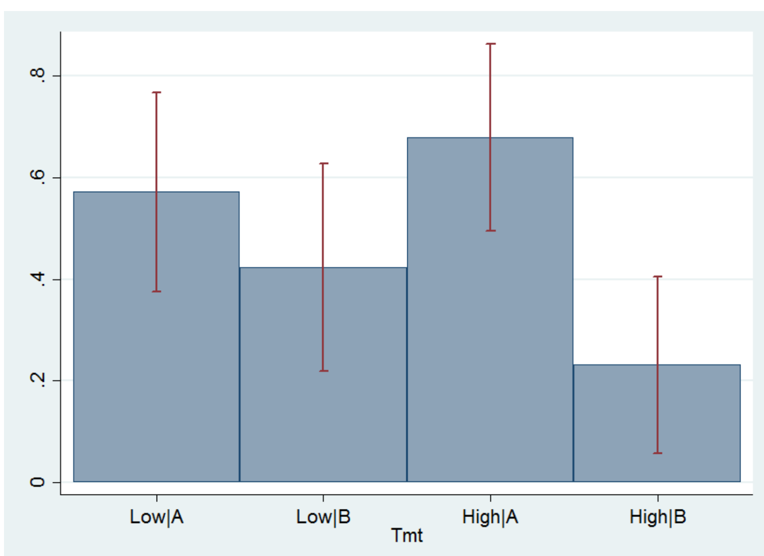


Figure 3: Asymmetry of in-group bias between HIGH and LOW (relative frequency of LEFT choices in the first decision phase by treatment; error bars show 95% confidence intervals).

The result is different for high-scoring subjects. In this case, 67.86% (19/28) of subjects chose LEFT in treatment High|A, and only 23.08% (6/26) did so in High|B (see the two rightmost bars in Figure 3). Therefore, the in-group bias amounts to 44.78 percentage points, which clearly differs from zero ($p = 0.001$). In comparison to low-scoring subjects, the difference-in-difference effect of 30 percentage points ($44.8 - 14.8$) is substantial and very close to the conventional level of significance ($p = 0.102$, compare last row in Table 2). We conclude that the in-group bias is strong and significant only for subjects who score high in performance.

In general, subjects' cognitive skills, and therefore their assignment to groups, might not be independent of social preferences (e.g., Ben-Ner et al., 2004) or other individual traits (e.g., Benjamin et al., 2013). Remember, however, that by differencing between conditions A and B, the design already accounts for this possibility.[22] Moreover, pooling the data across conditions A and B, the frequency of LEFT choices does not differ between high- and low-performing groups (0.46 (High) vs. 0.5 (Low), $p = 0.703$); that is, the groups show the same general extent of egalitarian behavior. This result is in line with Benjamin et al. (2013) who also find no correlation between cognitive ability and social preferences. Importantly, this finding already suggests that the asymmetry effect is unlikely to be a result of selection and unknown variables. The next subsection provides additional evidence supporting an interpretation of the asymmetry effect as causal.

4.2. In-group bias and individual performance in minimal groups

In the treatments involving minimal groups, 43.75% of 32 subjects in the role of I_A chose LEFT in treatment Min|A, compared to 20.83% of 24 subjects in Min|B. Accordingly, the in-group bias is 22.9 percentage points for minimal groups. This number is significantly different from zero at the 10% level ($p = 0.073$), and lies just between the 14.8 and 44.8 percentage points observed for real-group treatments. Therefore, we replicate earlier findings in the literature, according to which minimal groups already generate in-group bias. Furthermore, there is no asymmetry effect with assignment based on minimal groups.[23]

Figure 4 compares the behavior of high-scoring subjects (left panel) and low-scoring subjects (right panel) between the minimal-group treatments and the respective performance-based treatments.[24] The results show that the in-group bias of high-scoring subjects is smaller when the group assignment is "minimal", compared to "performance-based:" the bias is only 6.8 percentage points for high-scoring subjects in Min|A and Min|B (see the bars labeled High/Min|A and High/Min|B: 25.00% vs. 18.18%, $p = 0.692$); by contrast, we have already observed that the in-group bias is 44.8 percentage points for high-scoring subjects in performance-based groups (see the bars labeled High|A and High|B). For low-scoring subjects in minimal groups, the in-group bias is 31.9 percentage points (55.00% in Low/Min|A vs. 23.08% in Low/Min|B, $p = 0.070$); this compares to a bias of

[22]To illustrate, assume that the average high-scoring subject was more averse to inequality. This would show up as a difference in LEFT-choices between treatments HIGH and LOW, but the bias would difference out between conditions A and B.

[23]The in-group bias is 20.8% (37.5%-16.7%) in Kandinsky and 25% (50%-25%) in Klee subgroups; $p = 0.266$, χ^2 test, two sided.

[24]In the minimal group treatments, high-scoring subjects are those who perform above the median of all subjects in our experiment. Accordingly, there are 23 high-scoring and 33 low-scoring subjects in our minimal group treatments. Because of randomization, the distribution of types is orthogonal to treatment assignment; $p = 0.530$, χ^2 test, two sided.

14.8 percentage points in the LOW groups. Figure 4 reveals that most of the difference between the behavior of high and low-performing subject in minimal groups happens in condition A (compare the bars labeled High/Min|A and Low/Min|A): low-performing subjects choose LEFT more often than high-performing ones ($p = 0.098$ χ^2 test, two sided). The difference is slightly dampened in condition B, but the difference-in-difference effect is significant ($p = 0.084$, compare the second to last row in Table 2); this is, in minimal groups the in-group bias tends to be larger for subjects of type LOW than those of type HIGH.

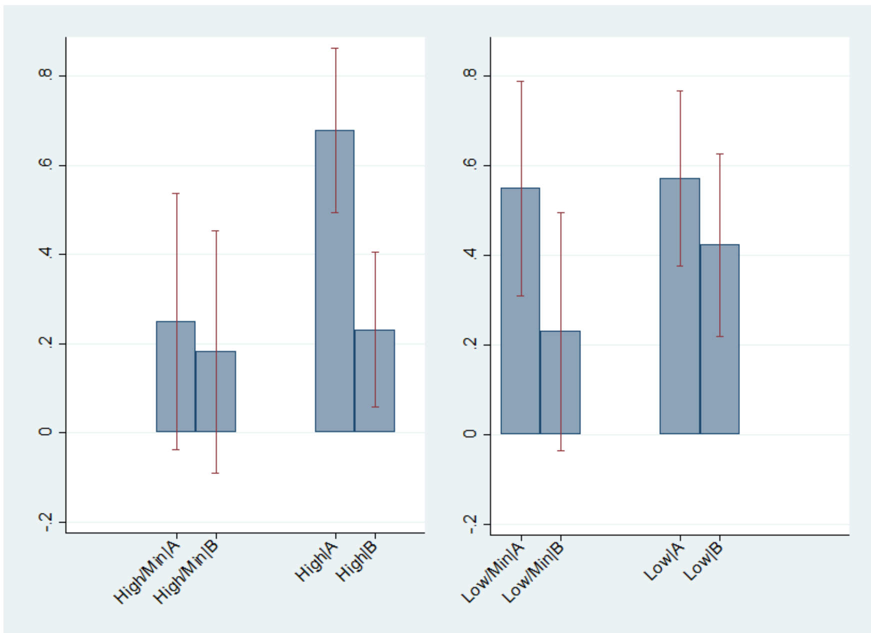


Figure 4: Frequency of LEFT choices split by treatment and performance (error bars show 95% confidence intervals).

Table 2 shows the results of a linear probability model summarizing our results.[25] The dependent variable takes a value of 1 if a subject chose LEFT, and zero otherwise. The regressions are split by subject type (high and low), they do not include a constant, and are estimated as follows: $LEFT_i = b_1 Min + b_2 Min \times A + c_1 High + c_2 High \times A + \epsilon_i$ (see column (1)), and $LEFT_i = b_1 Min + b_2 Min \times A + d_1 Low + d_2 Low \times A + \epsilon_i$ (column (2)), with Min, $High$, and Low indicating the respective treatment and A as dummy variable for being in condition A. Accordingly, 18.2% of high-scoring subjects choose LEFT in the role as member of a minimal group in condition B (see variable Min|B in column (1)). This number increases by 6.8 percentage points for subjects in minimal group in condition A (Min|A); that is, 6.8 percentage points is the estimate of the in-group bias among high-performing subjects in minimal groups. In comparison, column (2) repeats the estimation for subjects who score low in the test. Here, the estimated in-group bias is 31.9 percentage points in minimal groups (see Min|A). The difference-in-difference effect ($6.8 - 31.9 = -25.1$) is quantitatively

[25]The results are essentially unchanged if we use logistic regressions instead.

Table 2: Linear Probability Model: Dep.Var is LEFT

	(1) High	(2) Low
Min|B	0.182	0.231*
	(0.135)	(0.138)
Min|A	0.068	0.319*
	(0.187)	(0.177)
High|B	0.231**	
	(0.088)	
High|A	0.448***	
	(0.122)	
Low|B		0.423***
		(0.097)
Low|A		0.148
		(0.135)
Observations	77	87
R-squared	0.513	0.501
Min|A(1) = Min|A(2)	$p = 0.084$	
High|A(1) = Low|A(2)	$p = 0.102$	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

important and is significant at the 10 percent level (two-sided, see the row labeled "*Min|A(1) = Min|A(2)*"). This result shows that cognitive ability correlates with in-group bias in minimal groups.

The middle and lower parts of Table 2 replicate the findings in Figure 3. 23.1% of high-performing subjects choose LEFT in performance-based groups (see High|B in column(1)). The in-group bias is measured as the increase of this number between conditions A and B, and amounts to 44.8 percentage points (see High|A). For subjects who score low in performance-based groups, the in-group bias of 14.8 percent is not different from zero (see Low|A in column(2)). The differences-in-difference effect between high- and low-performing groups ($44.8-14.8 = +30$) is again quantitatively large; it very nearly significant at the 10 percent level (two-sided, see the row labeled "*High|A(1) = Low|A(2)*").

To summarize, our results suggest that performance in a test of cognitive ability influences the in-group bias at both the individual and the group level. However, the asymmetric in-group bias between groups strongly dominates and overturns the individual-level differences in minimal groups. This result indicates that the asymmetry effect is driven by what subjects know about the performance of the in-group relative to the out-group. It provides further evidence supporting the interpretation of the effect as causal.[26]

4.3. In-group punishment

The second decision phase is a one-to-one repetition of the first one with the exception that subjects in the role of passive in-group member I_P can now assign punishment points to the decision maker I_A. Subjects can assign up to 4 deduction points, each at a cost of 1 point for themselves, conditional on the decision maker choosing either RIGHT or LEFT. We define $P_R = \{0, 1, 2, 3, 4\}$ and $P_L = \{0, 1, 2, 3, 4\}$ as the decision to punish the decision RIGHT and LEFT, respectively.

Aggregated over all conditions, 23.2% (38 out of 164) of subjects punish the decision maker for choosing RIGHT; only 7.3% (12 out of 164) punish I_A for choosing LEFT. However, these numbers differ substantially between treatments. Table 3 shows the distribution of deduction points for each treatment split up by whether the decision maker has chosen RIGHT or LEFT.

Tmts	decision	Frequency of punishment					
		0 p.	1 p.	2 p.	3 p.	4 p.	Total p.
Min|A	P_R	23	2	3	2	2	22
	P_L	29	1	0	1	1	8
Min|B	P_R	24	0	0	0	0	0
	P_L	23	1	0	0	0	1
High|A	P_R	19	3	0	0	6	27
	P_L	27	0	0	0	1	4
High|B	P_R	24	1	0	0	1	5
	P_L	25	0	0	0	1	4
Low|A	P_R	19	2	2	3	2	23
	P_L	26	1	1	0	0	3
Low|B	P_R	17	5	3	1	0	14
	P_L	22	3	0	1	0	6
Total	P_R	126	13	8	6	11	91
	P_L	152	6	1	2	3	26

Table 3: Frequency of punishment per treatment and decision.

Table 3 shows that punishment is asymmetric in three dimensions. First, if we ignore the subtleties of some treatments for a moment, we observe that punishment is overwhelmingly directed toward decisions that implement the RIGHT distribution: averaged over all treatments, $P_R = 0.5549$ versus $P_L = 0.1585$, which is significantly different ($p < 0.0001$). This pattern mirrors the

[26] In all conditions, the decision makers are informed about their own absolute test scores; in addition, in performance-based groups, they learn something about their own scores relative to other subjects. It might be argued that this difference matters because decision makers are more likely to feel like strong (or weak) subjects. However, if we pool the data across conditions A and B, high-scoring decision makers are not more egoistic in minimal groups than in performance-based groups, $p = 0.420$ (regression-based t-test using LEFT as the dependent variable; independent variables are treatment (HIGH in comparison to MIN as the left-out category) and a control for the condition (A vs. B); the reported p-value refers to the estimated coefficient of the treatment dummy). Similarly, the pooled outcomes do not differ between minimal and low-performing groups ($p = 0.297$). Therefore, the data do not support the view that this argument is important in our design.

behavior in phase one. Choosing LEFT is costly to subjects, and therefore an expression of living up to a social norm. Thus, it might be expected that subjects who choose LEFT themselves are more likely to punish others for choosing RIGHT, that is, they are more willing to enforce the norm. In line with this interpretation, the extent and direction of punishment differs strongly between subjects depending on their choices in the role of I_A: $P_R = 0.9296$ for subjects who chose LEFT, versus $P_R = 0.2688$ for those who chose RIGHT in phase 1 ($p = 0.003$).

Second, there is substantially more punishment P_R in condition A than in B, both in MIN ($p = 0.005$) and HIGH ($p = 0.026$). Punishment for choosing LEFT (P_L) is not significantly different between conditions A and B for MIN, LOW, and HIGH groups ($p = 0.436$, $p = 0.349$ and $p = 0.958$, respectively). The behavior in these treatments is in line with, for example, the assumptions of Fehr and Schmidt (1999), according to which subjects dislike inequality, but they dislike unfavorable inequality (in condition A) more than favorable (in condition B).[27] If subjects anticipate the observed asymmetry of punishment between conditions A than B, this would contribute to further enforcing a norm of in-group loyalty. Our results indicate that such a mechanism might evolve in treatments MIN and HIGH.[28]

Third, subjects of type LOW stand out in that they punish their in-group peers for choosing the materially self-interested option (RIGHT) in condition B as well. In comparison to the other treatments, variable P_R takes a higher value in LOW|B than in MIN|B ($p = 0.004$). Between LOW|B and HIGH|B, the difference is less pronounced, but still significant, based on a one-sided test ($p = 0.070$). As a consequence, in groups with members of type LOW, the punishment behavior does not differ between conditions A and B ($p = 0.804$).[29]

In condition A, the punishment of RIGHT (P_R) signals a norm against unfavorable treatment of out-group members. One way to interpret why this behavior differs between treatments is that once they know that they are low performers, people favor a norm of equality (in low-performing groups, subjects are punished for not choosing the equal distribution both in Low|A and Low|B); otherwise, they favor a norm of entitlement (in high-performing groups, subjects are punished for not choosing the equal distribution in High|A but not in High|B). This interpretation fits with the results of Barr et al. (2015), who show that an individual's tendency to acknowledge earned entitlement is associated with his or her economic status relative to others. Note that this behavior is difficult to reconcile with simple inequality aversion. This is indicated, for instance, by the low

[27] Negative reciprocity would be another reason for punishing the decision maker for choosing RIGHT in condition A (I_P receives a payoff of 16 points rather than 20 points). However, it would then be expected that subjects also punish the decision maker for choosing LEFT in condition B (I_P receives a payoff of 20 points rather than 26 points); therefore, additional assumptions, like loss aversion, would be needed to rationalize the asymmetric pattern of punishment between conditions A and B.

[28] We do not report the choice behavior of phase 2 because it is very similar to phase 1. This finding might of course change in a design that permits a repeated number of phases with feedback on punishment after each phase. However, our findings seem to be in line with a study by Weng and Carlsson (2015), who find no effect of punishment on cooperation in teams with strong identity.

[29] Table A.1 in the Appendix further splits the data by individual performance in minimal groups: Low/Min|A, Low/Min|B, High/Min|A and High/Min|B. It shows that P_R also differs significantly between Low/Min|B and Low|B ($p = 0.031$). This indicates again that the asymmetric punishment pattern is unlikely a result of selection.

level of punishment P_R in the control treatment Min|B, which is 0.

Table 3 suggests that the observed effects predominantly come from reactions at the extensive margin (whether or not the subjects punish), rather than on the intensive margin (how much they punish). This impression is formally corroborated by a two-part hurdle model (Cragg, 1971), using P_R as the dependent variable (see, Table 4). The model accounts for the possibility of separating the decision to punish (selection model) from the decision of how much to punish (outcome model). We run the regression separately for conditions A and B. The omitted category is Min.

There are no treatment effects on the intensive margin (see the upper pane of Table 4). We observe significant effects only at the extensive margin (see lower pane of Table 4). These effects are in line with the results from above. First, subjects who have chosen LEFT themselves are more likely to punish the decision maker for choosing RIGHT (coef. 0.838 and 1.141 in condition A and B, respectively). Second, there is a positive and significant coefficient of Low only in condition B (coef. 1.510 in condition B), indicating that the decision maker is punished for choosing RIGHT more likely in treatment Low than in High.[30]

Remember that subjects in the role of I_P enter two punishment decisions, only one of which is relevant, conditional on whether I_A has chosen RIGHT or LEFT. So far, we only considered P_R in the regression, because only a few subjects punish the decision maker for choosing LEFT. Alternatively, we define $P_{R-L} = \{P_R - P_L\ , 0\}$ if I_P punishes RIGHT {*more than, equally to or less than*} LEFT. Table A.2 in the Appendix repeats the hurdle regression for P_{R-L}, and the results are qualitatively the same as in Table 4.

Finally, we asked the subjects in the role of O (the passive out-group members) to state their expectations about how many subjects in the role of I_P would punish I_A (see section 1). The subjects were paid for correct expectations and, other than this, they had no active part in the experiment. Therefore, expected punishment is an additional and independent source of data for testing for asymmetric identity norms between treatments. Table A.3 in the Appendix provides summary statistics and simple non-parametric tests for this data. Similar to before, we define $P^e_R = \{0, 1, ..., 10\}$ ($P^e_L = \{0, 1, ..., 10\}$) as the expected punishment for choosing RIGHT (LEFT), and we define $P^e_{R-L} = \{P^e_R - P^e_L\ , 0\}$ if O expects that I_P punishes RIGHT {*more than, equally to or less than*} LEFT.[31] Aside from the fact that subjects overestimate the actual extent of punishment, the expectations fully corroborate our findings from actual punishment behavior. Accordingly, the subjects expect no difference in punishment between Low|A and Low|B ($p = 0.940$); but they do expect significant differences in P^e_{R-L} both, between Min|A and Min|B ($p = 0.017$), and between High|A and High|B ($p = 0.043$) (see, Table A.3 in the Appendix).

[30] The omitted variable is Min|B. However, subjects happen not to punish in Min|B. Therefore, both High|B and Min|B jointly form the omitted category in the regression.

[31] P^e_L (P^e_R) are taken from the interval $\{0, 1, ..., 10\}$ because we ask subjects in the role of O about their expectation of how many out of 10 subjects in the role of I_P will punish LEFT (RIGHT).

Table 4: Punishment of the decision RIGHT

	Condition A	Condition B
Punishing RIGHT		
LEFT	-0.141	0.136
	(0.634)	(0.792)
High	0.646	
	(0.645)	
Low	0.143	-1.037
	(0.642)	(0.912)
Constant	2.447***	2.344**
	(0.621)	(1.081)
N	27	11
Selection		
LEFT	0.838***	1.141**
	(0.305)	(0.449)
High	0.016	
	(0.355)	
Low	0.056	1.510***
	(0.352)	(0.473)
Constant	-1.042***	-2.341***
	(0.301)	(0.482)
N	88	76
Wald-χ^2	9.234**	21.215***

Table notes. Hurdle models. Dependent variable: Punishment points $P_R = \{0, 1, 2, 3, 4\}$. $^*p \leq 0.1$, $^{**}p \leq 0.05$,$^{***}p \leq 0.01$.

4.4. Efficiency

In this section, we briefly discuss the potential effects of our results on efficiency, by looking at aggregate payoff differences between identity groups. We note that these results are unlikely to generalize, because they also depend on experimental design. In-group bias is defined in terms of variance across the in- and out-group dimension. Whether or not the asymmetric in-group bias translates into differences in means depends on further factors, such as selection effects (which were largely absent in our experiment) and the structure of decisions with regard to the consequences for ones' own and others' payoffs (which was balanced across conditions in our design).

Total payoffs are $20 + 20 + 20 = 60$ in LEFT and $26 + 26 + 16 = 68$ RIGHT. We have already seen that the behavior averaged over both conditions A and B does not differ between treatments. Consequently, total payoffs (as well as the payoffs per subgroups) do not differ between LOW, HIGH and MIN.[32] In the second decision stage with punishment, the picture remains essentially the same. Payoffs (pooled across conditions and including the cost of punishment) do not differ between HIGH, LOW, and MIN.[33]

5. Discussion and conclusion

We studied the role of performance differences in a task requiring cognitive effort in an experiment on in-group bias. We observe that (i) high-performing subjects exhibit no in-group bias as members of minimal groups, whereas low-performing subjects strongly do, and that (ii) groups consisting of subjects who score high show more in-group bias than those who score low.

Our first finding contributes to a recent literature showing that higher cognitive ability is correlated with less biased behavior (e.g., Benjamin et al., 2013). It suggests that the characteristics of subjects included in the sample are likely to be more important than previously thought in the literature on in-group bias. The result also helps us understand our second finding, suggesting that the direction of the asymmetric in-group bias is not simply a result of subjects of different ability reacting differently to being categorized in arbitrary groups.

The literature offers two conceptual explanations for our second result. One is common in social psychology and assumes that subjects derive a stronger sense of self from favorable group comparisons. The other explanation has been largely overlooked in the literature so far, and rests on a meritocratic notion of entitlements in the context of groups. Both explanations apply simultaneously, and are notoriously difficult to separate.

The asymmetry effect of performance perceptions between groups implies that members of low-status groups are less loyal to other members of their group. In terms of welfare, there are situations in which this kind of behavior could have large negative effects. For example, Gill and Stone (2015) show that entitlement considerations can mitigate the negative incentive effects within teams. Consequently, a lack of group loyalty might negatively affect cooperation incentives in low-performing groups.[34]

On the other hand, group loyalty is not always desirable. To illustrate this, consider the parameters we implemented in our experiment. In the conditions involving tension between group loyalty and self-interest, behavior that favors the in-group member decreases the overall sum of payments. This effect holds because, alongside avoiding negative consequences for the in-group member, the group-loyal choice is worse for the out-group member. In-group loyalty might therefore be important in perpetuating social inequality, for example when it gives rise to group conflict or when it hinders skills being used optimally owing to out-group discrimination (see Chowdhury

[32] In the first decision stage without punishment, the average payoffs are 21.71, 21.09, and 21.19 in MIN, HIGH, and LOW, respectively; p-values from a χ^2-test (two-sided) are as follows: LOW vs. HIGH $p = 0.878$, LOW vs. MIN $p = 0.427$, and HIGH vs. MIN $p = 0.344$.

[33] The average payoffs are 21.5, 20.59, and 20.67 in MIN, HIGH, and LOW, respectively: LOW vs. HIGH $p = 0.975$, LOW vs. MIN $p = 0.171$, and HIGH vs. MIN $p = 0.214$.

[34] Similarly, research in social psychology argues that the lack of group identity might have a negative effect on work motivation and performance. See van Knippenberg and Ellemers (2003)

et al., 2016; Bandiera et al., 2009). Another example in which group loyalty leads to bad outcomes is given by Hadnes et al. (2013). In a study of entrepreneurial activity in Africa, these authors show that group-sharing norms lead to substantial inefficiencies because they reduce individual incentives to provide effort.

Economic success often is correlated with the ability and willingness to exert effort, for example, in high-quality jobs. From this perspective, our findings shed new light on elite behavior. Sokoloff and Engerman (2000), Acemoglu et al. (2005), and Glaeser et al. (2003) highlight the importance of elite behavior for the development of countries. As pointed out by Paetzel and Traub (2017) and Côté et al. (2015), elite behavior depends crucially on the social distance between the elite and the remaining society. Paetzel and Traub (2017) formalize this idea in a skewness-adjusted social preference functional and provide some experimental evidence, whereas, Côté et al. (2015) provide macro-empirical evidence. Our findings suggest that social distance may depend on perceived performance and explain why high-performing groups (elites) discriminate against low-performing out-groups.

The reasons that groups are economically segregated are, of course, manifold. One of those reasons, which we did not consider in our design, is homophily (Currarini et al., 2009), which means that people tend to form groups with others who are similar. Further research may clarify whether the factors that lead to the formation of natural groups would mitigate or enhance the asymmetry effect. Based on the results of our experiment, it could be surmised that groups are equally biased as long as the group assignment is perceived to be caused by factors beyond individual control. On the other hand, a belief that the characteristics of the groups are determined by the effort of its members would suffice to generate the effects we describe in the experiment. Furthermore, psychological research provides evidence that subjective performance perceptions tend to be biased. On the one hand, successful people often downplay the role of luck as a reason for success — a phenomenon known as "illusion of control" (Langer, 1975). On the other hand, members of socially disfavored groups often perceive themselves as less than deserving — a phenomenon sometimes referred to as "system justification" (see Major, 1994; Jost, 2001). Such phenomena would further enforce the phenomenon we observed in our experiment.

Online Appendix (web supplement)

Supplementary material to this article can be found online at LINK.

References

Abeler, J., Marklein, F., 2016. Fungibility, labels, and consumption. Journal of the European Economic Association 15, 99–127.

Aberson, C.L., Healy, M., Romero, V., 2000. Ingroup bias and self-esteem: A meta-analysis. Personality and Social Psychology Review 4, 157–173.

Acemoglu, D., Johnson, S., Robinson, J.A., 2005. Institutions as a fundamental cause of long-run growth. Handbook of Economic Growth 1, 385–472.

Akerlof, G.A., Kranton, R.E., 2000. Economics and identity. Quarterly Journal of Economics 115, 715–753.

Alesina, A., Angeletos, G.M., 2005. Fairness and redistribution. American Economic Review 95, 960–980.

Alesina, A., Baqir, R., Easterly, W., 1999. Public goods and ethnic divisions. Quarterly Journal of Economics 114, 1243–1284.

Alesina, A., Giuliano, P., 2011. Preferences for redistribution, in: Benhabid, J., Bisin, A., Jackson, M.O. (Eds.), Handbook of Social Economics. Elsevier, North Holland. chapter 1, pp. 93–131.

Babcock, L., Loewenstein, G., Issacharoff, S., Camerer, C., 1995. Biased judgments of fairness in bargaining. American Economic Review 85, 1337–1343.

Ball, S., Eckel, C., Grossman, P., Zame, W., 2001. Status in markets. Quarterly Journal of Economics 116, 161–188.

Ball, S., Eckel, C.C., 1998. The economic value of status. Journal of Socio-Economics 27, 495–514.

Balliet, D., Wu, J., de Dreu, C., 2014. In-group favoritism in cooperation: A meta-analysis. Psychological Bulletin 140, 1556–1581.

Bandiera, O., Barankay, I., Rasul, I., 2009. Social connections and incentives in the workplace: Evidence from personnel data. Econometrica 77, 1047–1094.

Barr, A., Burns, J., Miller, L., Shaw, I., 2015. Economic status and acknowledgement of earned entitlement. Journal of Economic Behavior & Organization 118, 55–68.

Ben-Ner, A., Kong, F., Putterman, L., 2004. Share and share alike? gender-pairing, personality, and cognitive ability as determinants of giving. Journal of Economic Psychology 25, 581–589.

Benjamin, D.J., Brown, S.A., Shapiro, J.M., 2013. Who is behavioral? cognitive ability and anomalous preferences. Journal of the European Economic Association 11, 1231–1255.

Benjamin, D.J., Choi, J.J., Strickland, A.J., et al., 2010. Social identity and preferences. American Economic Review 100, 1913–1928.

Bernhard, H., Fischbacher, U., Fehr, E., 2006. Parochial altruism in humans. Nature 442, 912–915.

Bertrand, M., Mullainathan, S., 2001. Do people mean what they say? implications for subjective survey data. American Economic Review 91, 67–72.

Bettencourt, B., Charlton, K., Dorr, N., Hume, D.L., 2001. Status differences and in-group bias: meta-analytic examination of the effects of status stability, status legitimacy, and group permeability. Psychological Bulletin 127, 520.

Brandts, J., Charness, G., 2011. The strategy versus the direct-response method: A first survey of experimental comparisons. Experimental Economics 14, 375–398.

Brauer, M., 2001. Intergroup perception in the social context: The effects of social status and group membership on perceived out-group homogeneity and ethnocentrism. Journal of Experimental Social Psychology 37, 15–31.

Brewer, M.B., Manzi, J.M., Shaw, J.S., 1993. In-group identification as a function of depersonalization, distinctiveness, and status. Psychological Science 4, 88–92.

Burks, S.V., Carpenter, J.P., Goette, L., Rustichini, A., 2009. Cognitive skills affect economic preferences, strategic behavior, and job attachment. Proceedings of the National Academy of Sciences 106, 7745–7750.

Butler, J.V., 2014. Trust, truth, status and identity: An experimental inquiry. BE Journal of Theoretical Economics 14, 293–338.

Cappelen, A., Moene, K., Sørensen, E., Tungodden, B., 2013. Needs versus entitlementsan international fairness experiment. Journal of the European Economic Association 11, 574–598.

Cappelen, A.W., Hole, A.D., Sørensen, E.Ø., Tungodden, B., 2007. The pluralism of fairness ideals: An experimental approach. American Economic Review 97, 818–827.

Charness, G., Haruvy, E., Sonsino, D., 2007. Social distance and reciprocity: An internet experiment. Journal of Economic Behavior & Organization 63, 88–103.

Chen, Y., Li, S.X., 2009. Group identity and social preferences. American Economic Review 99, 431–457.

Cheng, J.T., Weidman, A.C., Tracy, J.L., 2014. The assessment of social status: A review of measures and experimental manipulations. The Psychology of Social Status, Cheng, J. T., Tracy, J. L., Anderson, C., eds. , 347–362.

Chowdhury, S.M., Jeon, J.Y., Ramalingam, A., 2016. Identity and group conflict. European Economic Review 90, 107–121.

Cohn, A., Maréchal, M.A., 2016. Priming in economics. Current Opinion in Psychology 12, 17–21.

Côté, S., House, J., Willer, R., 2015. High economic inequality leads higher-income individuals to be less generous. Proceedings of the National Academy of Sciences 112, 15838–15843.

Cragg, J.G., 1971. Some statistical models for limited dependent variables with application to the demand for durable goods. Econometrica 39, 829–844.

Croson, R., Gneezy, U., 2009. Gender differences in preferences. Journal of Economic Literature 47, 448–474.

Croson, R., Konow, J., 2009. Social preferences and moral biases. Journal of Economic Behavior & Organization 69, pp. 201–212.

Currarini, S., Jackson, M.O., Pin, P., 2009. An economic model of friendship: Homophily, minorities, and segregation. Econometrica 77, 1003–1045.

Ellemers, N., Kortekaas, P., Ouwerkerk, J., 1999. Self-categorisation, commitment to the group and group self-esteem as related but distinct aspects of social identity. European Journal of Social Psychology 29, pp. 371–389.

von Essen, E., Ranehill, E., 2013. Punishment and status. EFI Working Paper Series in Economics and Finance .

Falk, A., 2017. Status inequality, moral disengagement and violence. DIW Discussion Papers .

Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. Quarterly Journal of Economics , 817–868.

Fischbacher, U., 2007. z-tree: Zurich toolbox for ready-made economic experiments. Experimental Economics 10, 171–178.

Fiske, S.T., Cuddy, A.J., Glick, P., Xu, J., 2002. A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. Journal of Personality and Social Psychology 82, 878.

Fong, C., 2001. Social preferences, self-interest, and the demand for redistribution. Journal of Public Economics 82, 225–246.

Fowler, J.H., Kam, C.D., 2007. Beyond the self: Social identity, altruism, and political participation. Journal of Politics 69, 813–827.

Gill, D., Prowse, V., 2016. Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis. Journal of Political Economy 124, 1619–1676.

Gill, D., Stone, R., 2010. Fairness and desert in tournaments. Games and Economic Behavior 69, 346–364.

Gill, D., Stone, R., 2015. Desert and inequity aversion in teams. Journal of Public Economics 123, 42–54.

Glaeser, E., Scheinkman, J., Shleifer, A., 2003. The injustice of inequality. Journal of Monetary Economics 50, 199–222.

Goette, L., Huffman, D., Meier, S., 2012. The impact of social ties on group interactions: Evidence from minimal groups and randomly assigned real groups. American Economic Journal: Microeconomics 4, 101–115.

Greiner, B., 2015. Subject pool recruitment procedures: Organizing experiments with orsee. Journal of the Economic Science Association 1, 114–125.

Hadnes, M., Vollan, B., Kosfeld, M., 2013. The dark side of solidarity. mimeo. Frankfurt: Goethe University .

Hanaki, N., Akiyama, E., Funaki, Y., Ishikawa, R., 2015. Diversity in cognitive ability enlarges mispricing. GREDEG Working Papers Series .

Hargreaves Heap, S., Varoufakis, Y., 2002. Some experimental evidence on the evolution of discrimination, co-operation and perceptions of fairness. Economic Journal 112, 679–703.

Hargreaves Heap, S.P., Zizzo, D.J., 2009. The value of groups. American Economic Review 99, 295–323.

Harris, D., Herrmann, B., Kontoleon, A., Newton, J., 2015. Is it a norm to favour your own group? Experimental Economics 18, 491–521.

Hays, N.A., Blader, S.L., 2017. To give or not to give? interactive effects of status and legitimacy on generosity. Journal of Personality and Social Psychology 112, 17–39.

Heckman, J.J., Stixrud, J., Urzua, S., 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. Journal of Labor Economics 24, 411–482.

Heffetz, O., Frank, R.H., 2008. Preferences for status: Evidence and economic implications, in: Benhabid, J., Bisin, A., Jackson, M.O. (Eds.), Handbook of Social Economics. Elsevier, North Holland. chapter 1, pp. 69–91.

Henrich, J., Gil-White, F.J., 2001. The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. Evolution and Human Behavior 22, 165–196.

Hett, F., Kröll, M., Mechtel, M., 2016. Choosing who you are: The structure and behavioral effects of revealed identification preferences. SSRN working paper .

Hewstone, M., Rubin, M., Willis, H., 2002. Intergroup bias. Annual Review of Psychology 53, 575–604.

Hoff, K., Kshetramade, M., Fehr, E., 2011. Caste and punishment: The legacy of caste culture in norm enforcement. Economic Journal 121.

Hoffman, E., McCabe, K., Shachat, K., Smith, V., 1994. Preferences, property rights, and anonymity in bargaining games. Games and Economic Behavior 7, 346–380.

Hoffman, E., Spitzer, M.L., 1985. Entitlements, rights, and fairness: An experimental examination of subjects' concepts of distributive justice. Journal of Legal Studies 14, 259–297.

Holm, J., 2016. A model of redistribution under social identification in heterogeneous federations. Journal of Public Economics 143, 39–48.

Homans, G.C., 1974. Social behavior: Its elementary forms. North Holland.

Hoppe, E.I., Kusterer, D.J., 2011. Behavioral biases and cognitive reflection. Economics Letters 110, 97–100.

Jackson, M., 2011. An overview of social networks and economic applications, in: Benhabib, J., Bisin, A., Jackson, M. (Eds.), Handbook of Social Economics. North-Holland. volume 1, pp. 511 – 585.

Jost, J.T., 2001. Outgroup favoritism and the theory of system justification: A paradigm for investigating the effects of socioeconomic success on stereotype content. Cognitive Social Psychology: The Princeton symposium on the legacy and future of social cognition , 89–102.

Kamei, K., Putterman, L., Tyran, J.R., 2015. State or nature? endogenous formal versus informal sanctions in the voluntary provision of public goods. Experimental Economics 18, 38–65.

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

van Knippenberg, D., Ellemers, N., 2003. Social identity and group performance: Identification as the key to group-oriented efforts, in: Haslam, S., van Knippenberg, D., Platow, M., Ellemers, N. (Eds.), Social Identity at Work: Developing Theory for Organizational Practice. Psychology Press, North Holland, pp. 29–42.

Konow, J., 2000. Fair shares: Accountability and cognitive dissonance in allocation decisions. American Economic Review 90, 1072–1091.

Konow, J., 2003. Which is the fairest one of all? a positive analysis of justice theories. Journal of Economic Literature 41, 1188–1239.

Kranton, R., Pease, M., Sanders, S., Huettel, S., 2016. Groupy and non-groupy behavior: Deconstructing bias in social preferences. Mimeo .

Kranton, R.E., Sanders, S.G., 2017. Groupy versus non-groupy social preferences: Personality, region, and political party. American Economic Review 107, 65–69.

Krawczyk, M., 2010. A glimpse through the veil of ignorance: Equality of opportunity and support for redistribution. Journal of Public Economics 94, 131–141.

Lane, T., 2016. Discrimination in the laboratory: A meta-analysis of economics experiments. European Economic Review 90, 375–402.

Langer, E.J., 1975. The illusion of control. Journal of Personality and Social Psychology 32, 311.

Levin, S., Federico, C.M., Sidanius, J., Rabinowitz, J.L., 2002. Social dominance orientation and intergroup bias: The legitimation of favoritism for high-status groups. Personality and Social Psychology Bulletin 28, 144–157.

Lindqvist, E., Östling, R., 2013. Identity and redistribution. Public Choice 155, 469–491.

Luttmer, E.F., 2005. Neighbors as negatives: Relative earnings and well-being. Quarterly Journal of Economics 120, 963–1002.

Major, B., 1994. From social inequality to personal entitlement: The role of social comparisons, legitimacy appraisals, and group membership. Advances in Experimental Social Psychology 26, 293–355.

Messick, D.M., Sentis, K., 1983. Fairness, preference, and fairness biases, in: Messick, D.M., S., C.K. (Eds.), Equity Theory: Psychological and Sociological Perspectives. Praeger, New York NY, pp. 94–122.

Miller, D., 1999. Principles of Social Justice. Harvard University Press.

Mobius, M., Rosenblat, T., Wang, Q., 2016. Ethnic discrimination: Evidence from china. European Economic Review 90, 165–177.

Mollerstrom, J., Reme, B.A., Sørensen, E.Ø., 2015. Luck, choice and responsibilityan experimental study of fairness views. Journal of Public Economics 131, 33–40.

Mullen, B., Brown, R., Smith, C., 1992. Ingroup bias as a function of salience, relevance, and status: An integration. European Journal of Social Psychology 22, 103–122.

Newheiser, A.K., Olson, K.R., 2012. White and black american children's implicit intergroup bias. Journal of Experimental Social Psychology 48, 264–270.

Nicklisch, A., Paetzel, F., 2018. Need-based justice and distribution procedures: The perspective of economics, in: Kittel, B., Traub, S. (Eds.), Need-based Distributive Justice: An Interdisciplinary Perspective. Springer.

Ortmann, A., Hertwig, R., 2002. The costs of deception: Evidence from psychology. Experimental Economics 5, 111–131.

Paetzel, F., Traub, S., 2017. Skewness-adjusted social preferences: Experimental evidence on the relation between inequality, elite behavior, and economic efficiency. Journal of Behavioral and Experimental Economics 68, 130–139.

Piketty, T., 1998. Self-fulfilling beliefs about social status. Journal of Public Economics 70, 115–132.

Proto, E., Rustichini, A., Sofianos, A., 2016. Intelligence, personality and gains from cooperation in repeated interactions. CESifo Working Papers Series .

Putterman, L., Tyran, J.R., Kamei, K., 2011. Public goods and voting on formal sanction schemes. Journal of Public Economics 95, 1213–1222.

Rodriguez-Lara, I., Moreno-Garrido, L., 2012. Self-interest and fairness: Self-serving choices of justice principles. Experimental Economics 15, 158–175.

Rudman, L.A., Feinberg, J., Fairchild, K., 2002. Minority members' implicit attitudes: Automatic ingroup bias as a function of group status. Social Cognition 20, 294–320.

Rustichini, A., 2015. The role of intelligence in economic decision making. Current Opinion in Behavioral Sciences 5, 32–36.

Schmader, T., Johns, M., Forbes, C., 2008. An integrated process model of stereotype threat effects on performance. Psychological Review 115, 336–356.

Schniter, E., Shields, T.W., 2014. Ageism, honesty, and trust. Journal of Behavioral and Experimental Economics 51, 19–29.

Segal, C., 2012. Working when no one is watching: Motivation, test scores, and economic success. Management Science 58, 1438–1457.

Selten, R., 1988. The equity principle in economic behavior, in: Leinfellner, W., Eberlein, G. (Eds.), Models of Strategic Rationality. Springer, pp. 269–281.

de Séréville, C., Myers, B., 1994. Nøglen til succes i alle IQ-tester. Nørhavn A/S, Viborg.

Shayo, M., 2009. A model of social identity with an application to political economy: Nation, class, and redistribution. American Political Science Review 103, 147–174.

Sokoloff, K.L., Engerman, S.L., 2000. History lessons: Institutions, factors endowments, and paths of development in the new world. Journal of Economic Perspectives 14, 217–232.

Tajfel, H., 1970. Experiments in intergroup discrimination. Scientific American 223, 96–103.

Tajfel, H., 1982. Social psychology of intergroup relations. Annual Review of Psychology 33, 1–39.

Tajfel, H., Turner, J., 1979. An integrative theory of intergroup conflict. The Social Psychology of Intergroup Relations 33.

Tsutsui, K., Zizzo, D.J., 2014. Group status, minorities and trust. Experimental Economics 17, 215–244.

Turner, J.C., Hogg, M.A., Oakes, P.J., Reicher, S.D., Wetherell, M.S., 1987. Rediscovering the social group: A self-categorization theory. Basil Blackwell.

Tyler, T.R., 2006. Psychological perspectives on legitimacy and legitimation. Annual Review of Psychology 57, 375–400.

Weiss, Y., Fershtman, C., 1998. Social status and economic performance: A survey. European Economic Review 42, 801–820.

Weng, Q., Carlsson, F., 2015. Cooperation in teams: The role of identity, punishment, and endowment distribution. Journal of Public Economics 126, 25–38.

Online Appendix (web supplement)

Appendix A. Tables

Tmts	decision	Frequency of punishment					
		0 p.	1 p.	2 p.	3 p.	4 p.	Total p.
High|A	P_R	19	3	0	0	6	27
	P_L	27	0	0	0	1	4
High|B	P_R	24	1	0	0	1	5
	P_L	25	0	0	0	1	4
Low|A	P_R	19	2	2	3	2	23
	P_L	26	1	1	0	0	3
Low|B	P_R	17	5	3	1	0	14
	P_L	22	3	0	1	0	6
Min|A	P_R	23	2	3	2	2	22
	P_L	29	1	0	1	1	8
Min|B	P_R	24	0	0	0	0	0
	P_L	23	1	0	0	0	1
Low/Min|A	P_R	13	2	3	1	1	15
	P_L	17	1	0	1	1	8
Low/Min|B	P_R	13	0	0	0	0	0
	P_L	13	0	0	0	0	0
High/Min|A	P_R	10	0	0	1	1	7
	P_L	12	0	0	0	0	0
High/Min|B	P_R	11	0	0	0	0	0
	P_L	10	1	0	0	0	0

Table A.1: Frequency of punishment per treatment and decision.

Table A.2: Punishing RIGHT more than LEFT

	only A conditions	only B conditions
LEFT	-0.635	-0.126
	(0.710)	(0.892)
HIGH	0.489	
	(0.660)	
LOW	0.129	-1.496
	(0.686)	(0.952)
Constant	3.045***	2.613**
	(0.783)	(1.124)
N	24	9
Selection		
LEFT	1.426***	1.389***
	(0.297)	(0.439)
HIGH	0.205	
	(0.353)	
LOW	0.175	1.205***
	(0.354)	(0.468)
Constant	-1.787***	-2.577***
	(0.316)	(0.475)
N	88	76
Wald-χ^2	28.517***	21.598***

Table notes. Hurdle models. Dependent variable: Punishing RIGHT more than LEFT in points $[0, 1, 2, 3, 4]$. $*p \leq 0.1$, $**p \leq 0.05$,$***p \leq 0.01$.

Table A.3: Expectations of subjects in role O regarding I_P's punishment

Min|A ($N = 16$)	Min|B ($N = 12$)	
$P^e_{R-L} = 4.063$	$P^e_{R-L} = 1.75$	$p = 0.017$
Low|A ($N = 14$)	Low|B ($N = 13$)	
$P^e_{R-L} = 2.714$	$P^e_{R-L} = 2.462$	$p = 0.940$
High|A ($N = 14$)	High|B ($N = 13$)	
$P^e_{R-L} = 4.429$	$P^e_{R-L} = 2.308$	$p = 0.043$

Notes: Expectations of subjects in role O regarding I_P's punishment: $P^e_{R-L} = \{P^e_R - P^e_L\ ,0\}$ if O expects that I_P punishes I_A $\{$*more than, equally to or less than*$\}$ for choosing RIGHT than LEFT. $P^e_R = \{0, 1, ..., 9, 10\}$ ($P^e_L = \{0, 1, ..., 9, 10\}$) is defined as the expected punishment for choosing RIGHT (LEFT). p-values in parenthesis are based on two-sample Wilcoxon rank-sum (Mann–Whitney) tests.

Appendix B. Instructions

Instructions - part one

Welcome to the experiment. If you read the instructions carefully and follow the rules, you will have the opportunity to earn money. You will receive your payment in cash at the end of the experiment.In the experiment, we do not talk of Euros. Instead, all your payments are calculated in experimental points. The value of points is given by the following exchange rate:

$$1 \text{ point} = 0{,}35 \text{ Euro}.$$

During the experiment you are not allowed to speak to other participants. If you have any questions, please ask us, and we will answer your question in private. It is very important that you follow these rules. Otherwise, the results of this experiment have no value from a scientific perspective.

The experiment consists of **three parts**; every part is explained separately. The experiment will last approximately 60 minutes. We now explain the first part of the experiment.

Detailed information on the first part of the experiment

The first part of the experiment consists of one task *[two tasks in the MIN treatments]*, which are described as follows.

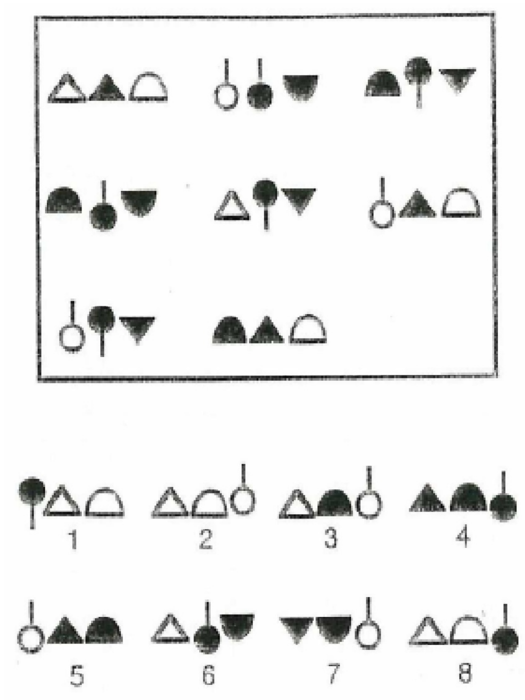
Task 1: You will observe 15 screens. On every screen you face a task. We ask you to solve as many tasks correctly as possible. There will be a time limit of 60 seconds per screen, otherwise the task counts as unsolved. At the end, you will be informed about how many tasks you have solved correctly. You will not receive any money for solving the tasks; nonetheless, we ask you to take this part seriously and try to solve as many tasks correctly as possible.

Task 2: *[This task is shown only to participants in the MIN treatments]* In the following, you will observe six screens in succession. On every screen, you will observe two paintings next to each other. One of the two paintings (you do not know which one) is always from Wassily Kandinsky and the other one is from Paul Klee. Your task is to indicate on each screen, which of the two paintings you like better.

Based on your decisions in task 2, you will be assigned a type.

- If you prefer the pictures by Kandinsky, the type KANDINSKY is assigned to you.
- If you prefer the pictures by Klee, the type KLEE is assigned to you.

Figure B.1: Sample Screen IQ-test



Note: Sample screen from the IQ-test. Participants have to find the correct symbol.

At the end of the first part, you will get to know whether you have been assigned type KANDINSKY or KLEE.

If you still have any questions, please raise your hand and wait quietly until one of the experimenters attends to you.

[The following feedback is provided to subjects at the end of the performance stage:]

Your score of correctly answered questions: Number♯

[The following information was only shown to participants in MIN treatments]
Based on your decisions, you are assigned into the group of participants preferring the following artist: [KLEE or KANDINSKY]

[The following information was only shown to participants in LOW and HIGH treatments]
In this experiment, the participants are split into two groups of equal size based on their scores of correctly answered questions.

- If you belong to the group of participants who correctly answered many questions relative to all participants you will be assigned type HIGH.
- If you belong to the group of the participants who correctly answered few questions relative to all participants you will be assigned type LOW.

Based on your relative performance in the task, you are assigned the following type: [LOW or HIGH]

Instructions — part two

You are now in the second part of the experiment. In this part, you and two other participants will form a group of three subjects. We call the participants in your group participants 1,2, and 3. Your role, regardsless whether you are participant 1, 2, or 3, is already determined and remains the same during the whole experiment. You will learn later in the experiment which role you are in.

Decision of participant 2

In this part of the experiment, only participant 2 will take a decision; participants 1 and 3 do not take decisions.

Consequences of the decision of participant 2

Participant 2 decides between two options. This decision has an impact on all participants in the group. The following Table B.4 shows the payment of participants 1, 2, and 3, depending on the decision of participant 2. If participant 2 decides on choosing LEFT (see left column), then all participants in the group receive 20 points. If participant 2 decides on RIGHT (see right column), then he or she receives 26 points, participant 1 receives 16 points, and participant 3 receives 26 points.

Table B.4: **Payoff of participants depending on the decision of participant 2**

	Participant 2 chooses left	Participant 2 chooses right
Participant 1	20	16
Participant 2	20	26
Participant 3	20	26

You will now see a decision screen on which participant 2 has to decide between LEFT and RIGHT. *[Between conditions A and B, the payoffs of participant 1 and 3 were swopped.]*

[Subjects are informed about their type (Kandinsky or Klee in minimal groups, and LOW or HIGH in performance-based groups)]

[Before they go to the decision screen, the two subjects who share the same type within the group are instructed on the screen about the decision mode:]

So far, you do not know whether you are in the role of participant 1 or 2. You will be informed about your actual role (participant 1 or participant 2) at the end of the experiment after you have reached a decision. To determine the payoffs, only the decision of the role of participant 2 is relevant. If you are in the role of participant 1, your decision does not affect the payoffs in the experiment.

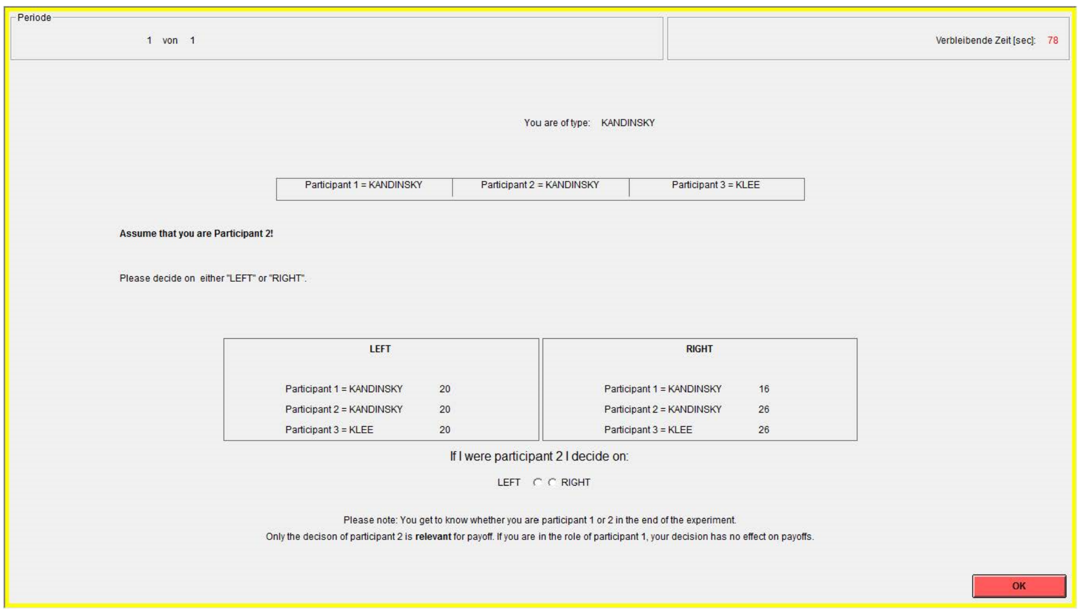


Figure B.2: Decision Screen.
Note: Sample screen from MIN treatment. In LOW and HIGH treatments, everything was exactly the same, but instead of KANDINSKY and KLEE, LOW and HIGH were shown to the participants.

Instructions — part three

Your payment of the second part is already determined and you will be informed about that payoff at the end of the experiment. Now, you receive the **instructions of the third part of the experiment:**

You are now in the third part of the experiment. This part is identical to the second part of the experiment, with the only difference being that now, participant 1 also takes a decision. In this part you are in the same group and in the same role as in the second part.

Decision of participant 2: The decision of participant 2 is identical to the decision in the second part of this experiment.

Decision of participant 1: Participant 1 has the opportunity to send “deduction points” to participant 2. Sending a deduction point is costly for participant 1 and receiving a deduction point is also costly for participant 2:

- For every deduction point that participant 1 sends to participant 2, participant 1 loses 1 point;
- For every deduction point that participant 2 receives from participant 1, participant 2 loses 3 points.

Participant 1 can send 4 deduction points at most.

- Example 1: Assume that participant 1 sends 3 deduction points. In this case, the payoff of participant 1 decreases by 3 points and the payoff of participant 2 decreases by 9 points (3 x 3 deduction points).
- Example 2: Assume that participant 1 sends no deduction point. In this case the payments of participants 1 and 2 remain unchanged.

In the following, you will observe a decision screen on which those in the role of participant 2 decide between LEFT and RIGHT. After that, another screen appears on which those in the role of participant 1 decide how many deduction points they want to send to those in the role of participant 2.

[Depending on the treatments, subjects have the label [Kandinsky or Klee in MIN and LOW or HIGH in treatments LOW and HIGH] and are assigned into treatment specific groups.]

After the third part, the experiment is over and you receive your payment in cash. If you have any questions, raise your hand and wait quietly until one of the experimenters attends to you.

Appendix C. Instructions in German

Instruktionen Teil 1

Willkommen zum Experiment. Wenn Sie die Instruktionen aufmerksam lesen und alle Regeln beachten, können Sie in diesem Experiment Geld verdienen. Das Geld wird im Anschluss sofort in bar an Sie ausbezahlt. Während des Experimentes sprechen wir nicht von Euro sondern von Punkten. Diese werden gemäß folgendem Wechselkurs umgerechnet:

$$1 \text{ Punkte} = 0{,}35 \text{ Euro.}$$

Während des gesamten Experiments ist das Sprechen mit anderen Teilnehmern nicht erlaubt. Wenn Sie Fragen haben, richten Sie diese bitte ausschließlich an uns. Wir beantworten Ihre Fragen gerne individuell. Die Einhaltung dieser Regel ist sehr wichtig. Andernfalls sind die Ergebnisse dieses Experimentes wissenschaftlich wertlos.

Dieses Experiment besteht aus **3 Teilen**, wobei jeder einzelne Teil nacheinander einzeln erläutert wird. Das Experiment wird voraussichtlich 60 Minuten dauern. Im Folgenden wird Ihnen nun der erste Teil des Experimentes erläutert.

Detaillierte Informationen zum 1. Teil des Experiments

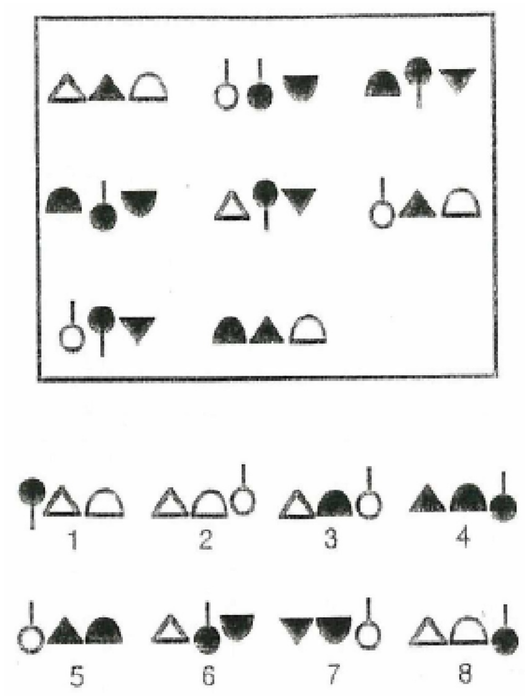
Der erste Teil des Experimentes besteht aus einer Aufgabe *[zwei Aufgaben in den MIN treatments]*, welche im Folgenden beschrieben werden.

Aufgabe 1: Sie werden hintereinander 15 Bildschirme sehen. Auf jedem Bildschirm sehen Sie eine Aufgabe. Wir bitten Sie, so viele Aufgaben wie möglich richtig zu lösen. Sie haben für die Lösung einer Aufgabe maximal 60 Sekunden Zeit. Andernfalls gilt die Aufgabe als nicht gelöst. Am Ende des ersten Teiles erfahren Sie wie viele Aufgaben Sie richtig gelöst haben. Sie erhalten kein Geld für das Lösen der Aufgaben: wir bitten Sie aber dennoch die Aufgaben ernst zu nehmen und so gut Sie können zu lösen.

Aufgabe 2: *[Diese Aufgabe sehen nur die Teilnehmer in den MIN treatments]* Im Folgenden werden Sie hintereinander sechs Bildschirme sehen. Auf jedem Bildschirm sehen Sie zwei Gemälde nebeneinander. Eines der beiden Gemälde (Sie wissen nicht welches) ist dabei immer von Wassily Kandinsky und das andere von Paul Klee. Ihre Aufgabe besteht nun darin, auf jedem Bildschirm anzugeben, welches der beiden Gemälde Ihnen besser gefällt. Entscheiden Sie einfach nach Ihrem Geschmack.

Auf Basis Ihrer Entscheidungen werden Sie dann einem Typ zugeordnet.

Figure C.1: Beispielbildschirm IQ-Test



Note: Beispielbildschirm IQ-Test. Teilnehmer müssen das passende Symbol finden.

- Wenn Ihnen die Bilder von Kandinsky relativ besser gefallen, wird Ihnen der Typ KANDINSKY zugeordnet.
- Wenn Ihnen die Bilder von Klee relativ besser gefallen, wird Ihnen der Typ KLEE zugeordnet.

Am Ende des 1. Teils des Experimentes werden Sie erfahren, ob Sie vom Typ KLEE oder KANDINSKY sind.

Falls Sie nun noch Fragen haben, heben Sie die Hand und warten Sie ruhig, bis jemand zu Ihnen kommt.

[Die folgende Information wurde den Teilnehmern nach dem Leistungstest angezeigt:]

Ihre Anzahl an richtig gelösten Aufgaben♯

[Die folgende Information wurde nur den Teilnehmern in den MIN treatments angezeigt.]
Basierend auf Ihren Entscheidungen, wurden Sie in die Gruppe der Teilnehmer zugeteilt, die folgenden Maler bevorzugen: [KLEE or KANDINSKY]

[Die folgende Information wurde nur den Teilnehmern in den LOW und HIGH treatments angezeigt]
In diesem Experiment werden die Teilnehmer anhand der richtig beantworteten Fragen in zwei gleich große Gruppen einsortiert.

- Wenn Sie zur Hälfte der Teilnehmer gehören, die vergleichsweise viele Fragen richtig beantwortet hat, wird Ihnen der Typ HIGH zugeordnet.
- Wenn Sie zur Hälfte der Teilnehmer gehören, die vergleichsweise wenig Fragen richtig beantwortet hat, wird Ihnen der Typ LOW zugeordnet.

Durch Ihr relatives Abschneiden im Wissenstest wird Ihnen der folgende Typ zugeordnet: [LOW oder HIGH]

Instruktionen — Teil 2

Sie befinden sich nun im 2. Teil des Experimentes. In diesem Teil des Experimentes bilden Sie gemeinsam mit zwei anderen Teilnehmern eine 3er-Gruppe. Wir nennen die Teilnehmer in Ihrer Gruppe Teilnehmer 1, 2, und 3. Ihre Rolle (das heisst, ob Sie Teilnehmer 1, 2, oder 3 sind) steht bereits fest und bleibt während des gesamten Experiments unverändert. Sie erfahren Ihre Rolle jedoch erst später.

Entscheidung des Teilnehmers 2:
In diesem Teil des Experiments trifft ausschließlich Teilnehmer 2 eine Entscheidung; die Teilnehmer 1 und 3 treffen keine Entscheidung.

Auswirkung der Entscheidung des Teilnehmers 2

Teilnehmer 2 entscheidet zwischen zwei Alternativen. Diese Entscheidung hat Auswirkungen auf alle Teilnehmer in der Gruppe. Die folgende Tabelle C.5 zeigt die Auszahlungen der Teilnehmer 1, 2 und 3 in Abhängigkeit der Entscheidung von Teilnehmer 2. Entscheidet sich Teilnehmer 2 für "Links" (siehe linke Spalte) erhalten alle Teilnehmer in der Gruppe 20 Punkte. Entscheidet sich Teilnehmer 2 für "Rechts" (siehe rechte Spalte) erhält er oder sie 26 Punkte, Teilnehmer 1 erhält 16 Punkte und Teilnehmer 3 erhält 26 Punkte.

In Kürze erscheint ein Bildschirm am Computer, auf dem sich der Teilnehmer in der Rolle des Teilnehmers 2 zwischen "Links" und "Rechts" entscheidet. *[Die Situationen A und B unterscheiden sich dahingehend, dass die Auszahlungen für Teilnehmer 1 und 3 vertauscht sind.]*

[Teilnehmer werden informiert welcher Typ (KLEE oder KANDINSKY in MIN treatments oder HIGH und LOW in den HIGH, LOW treatments) Ihnen zugeordnet wurde]

Table C.5: **Auszahlungen der Teilnehmer in Abhängigkeit der Entscheidung von Teilnehmer 2**

	Teilnehmer 2 entscheidet sich für Links	Teilnehmer 2 entscheidet sich für Rechts
Teilnehmer 1	20	16
Teilnehmer 2	20	26
Teilnehmer 3	20	26

[Bevor die Teilnehmer 1 und 2 den Entscheidungsbildschirm sehen, werden Sie über das Entscheidungsverfahren [Strategiemethode] informiert:]

Noch wissen Sie nicht, ob Sie in der Rolle des Teilnehmers 1 oder 2 sind. Erst am Ende des Experimentes, nachdem Sie Ihre Entscheidung getroffen haben, werden Sie erfahren, in welcher Rolle (Teilnehmer 1 oder Teilnehmer 2) Sie tatsächlich sind. Für die Auszahlungen relevant ist nur die Entscheidung des Teilnehmers 2. Wenn Sie in der Rolle des Teilnehmers 1 sind, so hat Ihre Entscheidung keinen Einfluss.

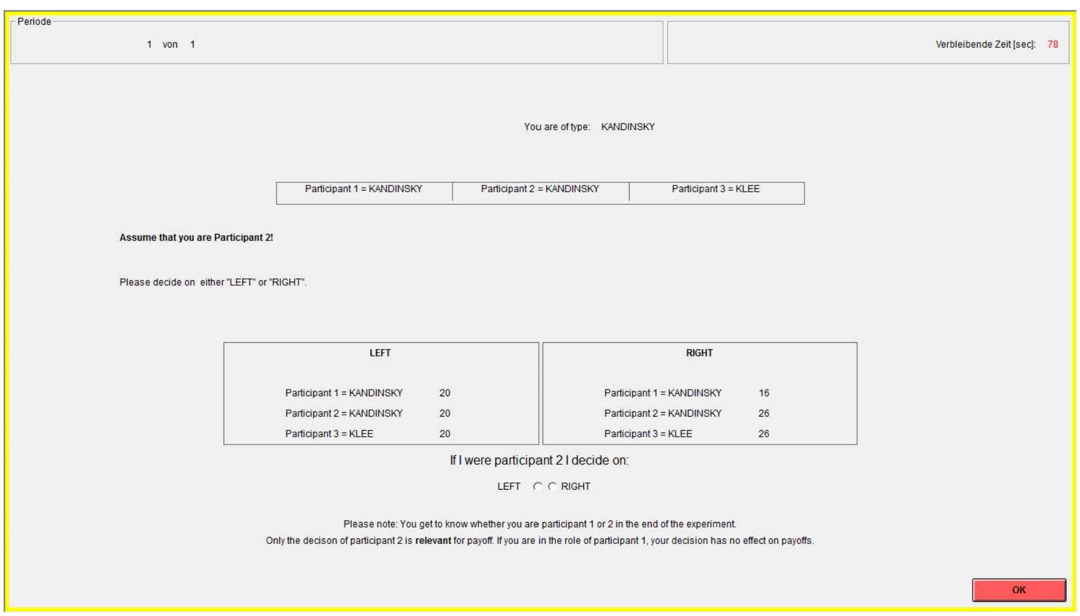


Figure C.2: Decision Screen.
Note: Sample screen from MIN treatment. In LOW and HIGH treatments, everything was exactly the same, but instead of KANDINSKY and KLEE, LOW and HIGH were shown to the participants.

Instruktionen — Teil 3

Ihre Auszahlung aus dem zweiten Teil des Experimentes steht fest und wird Ihnen am Ende des Experiments mitgeteilt. Sie erhalten nun die **Instruktionen zum dritten Teil des Experimentes:**

Sie befinden sich nun im 3. Teil des Experimentes. Dieser Teil ist identisch mit dem 2. Teil des Experimentes, mit der einzigen Ausnahme, dass nun auch Teilnehmer 1 eine Entscheidung trifft. Sie sind in diesem Teil des Experimentes in derselben 3er-Gruppe und in derselben Rolle wie zuvor.

Entscheidung des Teilnehmers 2 Die Entscheidung des Teilnehmers 2 ist identisch mit jener im 2. Teil des Experiments.

Entscheidung des Teilnehmers 1 Teilnehmer 1 hat die Möglichkeit, Teilnehmer 2 sogenannte "Abzugspunkte" zu senden. Ein Abzugspunkt kostet sowohl für Teilnehmer 1 als auch für Teilnehmer 2 Geld:

- für jeden Abzugspunkt, den Teilnehmer 1 an Teilnehmer 2 sendet, verliert Teilnehmer 1 einen Auszahlungspunkt;
- für jeden Abzugspunkt, den Teilnehmer 2 von Teilnehmer 1 empfängt, verliert Teilnehmer 2 drei Auszahlungspunkte.

Teilnehmer 1 kann höchstens 4 Abzugspunkte senden.

- Beispiel 1: Angenommen Teilnehmer 1 sendet drei Abzugspunkte. In diesem Fall reduziert sich die Auszahlung von Teilnehmer 1 um drei Punkte und jene des Teilnehmers 2 reduziert sich um neun Punkte (dreimal drei Abzugspunkte)..
- Beispiel 2: Angenommen Teilnehmer 1 sendet keine Abzugspunkte. In diesem Fall bleiben die Auszahlungen von Teilnehmer 1 und 2 unverändert.

In Kürze erscheint wieder ein Bildschirm am Computer, auf dem sich die Teilnehmer in der Rolle des Teilnehmers 2 zwischen "Links" und "Rechts" entscheiden. Zudem erscheint ein weiterer Bildschirm, auf dem sich die Teilnehmer in der Rolle des Teilnehmers 1 entscheiden, wie viele Abzugspunkte sie an Teilnehmer 2 senden wollen.

[Abhängig vom Treatment haben die Teilnehmer entweder die Typen KLEE und KANDINSKY oder HIGH und LOW erhalten und sind spezifisch nach Treatment in die entsprechende Gruppe eingeordnet worden.]

Nach diesem dritten Teil ist das Experiment zu Ende und Sie erhalten Ihre Auszahlung. Falls Sie nun noch Fragen haben, heben Sie die Hand und warten Sie ruhig, bis jemand zu Ihnen kommt.