The influence of social comparisons on cooperation and fairness

Inauguraldissertation zur

Erlangung des Doktorgrades der

Wirtschafts- und Sozialwissenschaftlichen Fakultät der

Universität zu Köln

2019

vorgelegt von

MA Jarid Emanuel Zimmermann

aus

Nelson, Kanada.

Referent:	Prof. Dr. Bettina Rockenbach
Korreferent:	Prof. Dr. Matthias Sutter
Tag der Promotion:	09.12.2019

Danke!

Ich danke Nora, Luc und meiner ganzen Familie, sowie allen, die dazu beigetragen haben dies möglich zu machen! Ganz besonderer Dank gilt Bettina Rockenbach und Robert Böhm, die mich bei all meinen Irrungen und Wirrungen stets unterstützt haben und es mir ermöglicht haben meinen Horizont zum Thema Kooperation und Fairness zu erweitern. Großer Dank geht auch an Matthias Sutter, der das Korreferat übernommen hat. David Buttelmann möchte ich für seine Unterstützung beim Kids-Experiment danken. Bei Caroline Stein, Anna Untertrifaller und Sebastian Butschek möchte ich mich für die mentale Unterstützung bei der Arbeit, aber auch für die Ablenkung davon, bedanken. Beim Team Rockenbach (Thomas Lauer, Heike Kirch, den beiden Sebastians, also Schneiders und Tonke, Lukas Wenner) und bei Stefania Bortollotti, Christoph Feldhaus, Felix Kölle, Matthias Praxmarer, Uta Schier und Claudia Zoller möchte ich mich für viele anregende Diskussionenen im Büro, in der Mensa und am Kaffeestand bedanken. Urs Fischbacher danke ich dafür, dass ich über mehrere Monate teil seiner Forschungsgruppe in Konstanz sein durfte. Aus Konstanz danke ich zudem Fabian Dvorak, der den Kontakt hergestellt hat, sowie David Grammling und Jan Hausfeld, die ihr Büro mit mir geteilt haben.

Contents

Chapter 1

Introduction

The ability to foster and maintain cooperation and fairness is a key determinant of (economic) well-being in human societies. When two or more individuals cooperate, they engage in joint actions that, even though costly to the individual, create mutual benefits (Bowles and Gintis, 2008). Similarly, fairness decisions require that an individual gives up a part of his or her own resources and re-distributes it to others (Kahneman et al., 1986). Both cooperation and fairness decisions have been shown to crucially depend on social information, that is, information about other people's behavior. Field studies have shown that people make higher contributions to public goods (Chen et al., 2010; Frey and Meier, 2004), are more likely to donate to charitable organizations (Shang and Croson, 2009), and refrain from littering (Cialdini et al., 1991), if others do so, too. Yet, some interventions have also produced undesirable effects (Schultz et al., 2007; Fellner et al., 2013; Bhanot, 2017), and the underlying psychological processes remain unclear. Hence, in a series of economic experiments, this dissertation sets out to investigate how and why social information impacts cooperation and fairness decisions.

Early advances in behavioral-economic theory have focused on people's prosocial motivations to explain reactions to social information (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Charness and Rabin, 2002). For example, they modeled prosocial motivations as the result of payoff-comparisons between direct interaction partners, i.e., 'social preferences' for equitable outcomes (Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999). From this perspective, social information affects decisions because it reveals information about others' resources and people dislike having less (or more) than others. While the early theories have substantial explanatory power, they fail to explain the 'flickery' nature of social behavior, e.g., why fair-minded people would actively seek (and follow) information suggesting unfair behavior. Importantly, the early

theories have abstracted from people's desire for positive signals about themselves, which are key in social-psychological theory (e.g., Leary and Baumeister, 2000) and so-called 'second-generation models' of social preferences (Bodner and Prelec, 2003; Bénabou and Tirole, 2006; Andreoni and Bernheim, 2009; Bénabou and Tirole, 2011). This dissertation thus presents four economic experiments focusing particularly on how people's desires to send positive signals about their own prosociality, i.e, self- and social signaling, help explain the impact of social information on cooperation and fairness decisions.

From a social psychological perspective, people look to social information to answer questions they find relevant for themselves: How should I evaluate my own actions? How will others evaluate my actions? Where do I stand relative to others? In other words, they use social information with the purpose of social comparison, i.e., to evaluate themselves and their decisions relative to their social environment and its social norms (Cialdini et al., 1991; Festinger, 1954; Taylor et al., 1996; Tesser, 1988).

Self-signaling motives, i.e., the desire to appear prosocial to oneself, might arise because people lack perfect knowledge about their 'true' motivations, so they need to rely on 'hard facts' such as their own actions to infer 'who they really are'. Equivalently, one could say that people want to see themselves as prosocially motivated persons or that they care about their self-image (Bem, 1972; Bénabou and Tirole, 2006; Leary and Baumeister, 2000; Mazar et al., 2008). In this view, people might behave more cooperatively due to social information because it makes salient their prosocial identity. Speaking to self-signaling motives, Chapter 2 explores the conditions under which people try to boost their self-image as a prosocial person by cooperating in their own team, if their team's cooperativeness is compared to that of another team (Tajfel and Turner, 1979; Turner, 1975). Relatedly, Chapter 5 investigates whether or not people's desire for prosocial identity motivates them to set an example for others.

Social-signaling motives, i.e., the desire to appear prosocial to others, might arise since people care about how spectators, be they real or imaginary, perceive their motivations (Andreoni and Bernheim, 2009; Ariely et al., 2009; Bénabou and Tirole, 2006). An external spectator cannot know an individual's preferences, so he or she needs to rely on an individual's actions to infer his or her 'true' motivations. For instance, people may desire to appear prosocial to others in order to (i) receive higher social approval (Leary, 2007) or (ii) to be chosen as partners (Barclay and Willer, 2007). Note that if the desire to send positive signals about oneself enters preferences, inferences about the own motivations are 'strategic' (Bénabou and Tirole, 2016; Kunda, 1990): Individuals might prefer positive over negative signals about their prosocial motivation. In

Chapter 3, we therefore test whether people acquire social comparison information strategically, if it might negatively affect not only their self-, but also their social-signaling. Building on Chapter 3, Chapter 4 studies when and how children develop the ability to strategically select social comparison information.

So far, we have implicitly assumed that cooperation and fairness decisions are (perceived as) socially desirable. Yet, what is perceived as socially desirable crucially depends on the prevalent social norms in society (Herrmann et al., 2008), and social information can crucially affect people's norm perceptions (Krupka and Weber, 2009). *Descriptive* norms depend on what 'is typically done' in a given situation; *injunctive* norms depend on what 'should be done' in a given situation (Cialdini et al., 1991). In conjunction with the own 'true' prosocial motivations, social norms will affect how self- and social signaling works by affecting which behaviors are considered as socially desirable and thus how people's behavior affects their appearance (Mazar et al., 2008). Accordingly, the impact of social information on normative perceptions will play an important, yet tacit, role in each of the following chapters.

1.1 Overview and Main Findings

The following section provides an overview of the economic experiments featured in this dissertation and summarizes their main findings.

Chapter 2, entitled "United we stand, divided we fall: The limitations of between-group comparisons for fostering within-group cooperation", is published as Böhm et al. (2018). We investigate how between-group comparisons affect within-group cooperation decisions. Previous work has shown that the provision of information about the average contributions of a payoff-independent comparison group can increase within-group contributions in the public goods game. This study extends previous work by studying how (the perception of) between-group asymmetries in the between-group comparison affects within-group contributions. In particular, one group is disadvantaged because it faces a lower marginal-per-capita-return (MPCR) to cooperation than its comparison group. To test how the *perception* of disadvantages in one group affects within-group cooperation, we vary whether or not group members are aware of the between-group asymmetry.

We find that participants generally desired to be 'ahead of' their comparison group, providing evidence for the importance of people's desire for positive (group) distinctiveness. Yet, in addition, the perception of between-group asymmetries led members of disadvantaged

groups to respond more sensitively to within-group comparisons, i.e., to how their own payoffs compared to those of their group members. More precisely, participants more strongly reduced their contributions when their team members were free-riding. Particularly those participants who more strongly perceived that contributions are not comparable between groups reduced their contributions more strongly. Our findings can be explained from a self-signaling perspective: When the information made salient their group's disadvantages, members of disadvantaged groups focused more strongly on within-group comparisons (e.g., payoff-comparisons with their team members) to maintain a positive self-image.

Chapter 3 is entitled "The driving forces behind information acquisition in social decisions". This chapter is joint work with Robert Böhm and Bettina Rockenbach. We let people choose which social comparison information they want to acquire (or avoid) before making a fairness decision – a transfer from their own endowment to an anonymous recipient in the dictator game. Previous work has shown that people avoid information about the adverse consequences of their own actions, if the information is available only to themselves. Contrary to previous work, in our experiment, we varied whether or not the acquired social information could become public or not, so the social information could not only affect what people think of themselves (self-signaling), but also what other people think of them (social-signaling).

We find that if the acquired information became public and thus more strongly influenced people's prosocial appearance, people were more likely to acquire social information suggesting low prosocial behavior, and they used it as an excuse for making a low transfer. We also provide clean evidence that those who acquired information suggesting low prosocial behavior were also less prosocially motivated *per se*. In particular, they transferred less than others, irrespective of the social information they received. This experiment thus shows how prosocial motivation and social signaling interact in information acquisition. From a policy perspective, our results suggest that making social information public may produce unintended results, if those with a low prosocial motivation use the information in a strategic way.

Chapter 4, "Social-Comparison Engineering in Children" is joint work with David Buttelmann, Robert Böhm und Bettina Rockenbach. Here we build on the finding from Chapter 3 that adults choose social comparison information strategically: We ask when and how the ability to choose social comparisons strategically, that is, even against the own preferences, develops in children. In a field study, children made drawings and, subsequently, they could choose between low-accuracy and high-accuracy drawings to serve as their comparison standard. In one treatment, children won a prize if their drawing was more

accurate than the comparison standard. Hence, children could win and enhance their appearance by choosing a low-accuracy drawing – making their own drawing appear more accurate. In another treatment, children won if their drawing was less accurate than the comparison standard. We also elicited children's basic preference in a treatment, where there was no evaluation at all.

Notably, children exhibited a strong basic preference for high-accuracy compared to low-accuracy drawings when there was no evaluation. As a result, children needed to overcome an inner conflict in the treatment where choosing the low-accuracy drawing made them win. In this situation, only children from 6-7 years of age, but not younger children, managed to strategically choose the comparison standard that made them win. By contrast, children of all ages managed to win similarly often if choosing the high-accuracy drawing made them win. Children's developing abilities to inhibit conflicting impulses, but not their desire to win, or their desire to be liked by others, could explain the age differences in strategically choosing social comparisons. This suggests that children's ability to inhibit conflicting impulses is not only important for individual decisions, such as resisting temptations, but also for success in social situations more generally.

In Chapter 5, entitled "The desire to set an example: Do people cooperate more if they are influential?", I investigate when and how people desire to set an example for others. Compared to previous studies, this study aims to isolate people's psychological motives to set an example for others by focusing on situations where there are no monetary incentives to set an example. Participants ('first-team players') decided on a contribution to their team's public good, and, subsequently, their contributions provided information for 'follower-teams' that independently played their own respective public goods games. To vary whether or not first-team players had social influence via the contribution information they provided, I varied whether first-team players' decisions were revealed to follower-teams *before* or *after* follower-teams made their decisions. To vary the salience of social norms, participants either had the opportunity to write a recommendation about what follower-teams should contribute or did not have such an opportunity.

I find that only when first-team players had social influence and could write a recommendation, first-team players reported a higher motivation to set an example and also made higher contributions compared to first-team players who did not have social influence and could (or could not) write a recommendation. Hence, only when the opportunity to write a recommendation made salient injunctive norms, people's self-image as a prosocial person became sufficiently salient so that people desired to set an example. As a side-finding, I further

document follower-team player's intriguing inclination to follow high (normative) but not low (counter-normative) contributions.

1.2 Co-Authors and Acknowledgements

This dissertation was financed by the grant "The influence of social comparisons on cooperation and fairness" from the German Research Foundation [grant number 246329797] which was awarded to Bettina Rockenbach as a part of the research group "Relativity in Social Cognition" (speakers Thomas Mussweiler and Christian Unkelbach).

Chapter 2 was co-authored by Robert Böhm and Bettina Rockenbach and is published as Böhm et al. (2018). The authors gratefully acknowledge funding from the German Research Foundation [grant number 246329797]. We are greatly indebted to fruitful comments from and discussion with Thomas Mussweiler, Wilhelm Hofmann, Christian Unkelbach, and other seminar participants of the research unit on "Relativity in Social Cognition".

Chapter 3 was co-authored by Robert Böhm and Bettina Rockenbach and is based on Böhm et al. (2019). The authors gratefully acknowledge funding from the German Research Foundation [grant number 246329797]. We are greatly indebted to fruitful comments from and discussion with Thomas Mussweiler, Wilhelm Hofmann, Christian Unkelbach, and other seminar participants of the research unit on "Relativity in Social Cognition". We also thank seminar participants at the conference of the 2018 ESA European Meeting and the *theem* conference (2017).

Chapter 4 was co-authored by Robert Böhm, David Buttelmann and Bettina Rockenbach and is based on Zimmermann et al. (2019). The authors gratefully acknowledge funding from the German Research Foundation [grant number 246329797]. We are greatly indebted to fruitful comments from and discussion with Wilhelm Hofmann, Christian Unkelbach, and other seminar participants of the research unit on "Relativity in Social Cognition" as well as participants of the Cologne Social Cognition Meeting 2018, the GSDS Conference in Konstanz, and the 2019 ESA Asia-Pacific Meeting.

Chapter 5 is based on Zimmermann (2019). The author gratefully acknowledges funding from the German Research Foundation [grant number 246329797] and from the Junior Start-Up Grant "Does the prospect of being influential increase cooperation?" of the Center for Social and Economic Behavior at the University of Cologne. I am greatly indebted to comments on earlier drafts of this paper from Caroline Stein, Robert Böhm, and Thomas Lauer.

Chapter 2

United we stand, divided we fall

The limitations of between-group comparisons for fostering within-group cooperation

2.1 Introduction

Competition between work teams is a common means to increase team performance. Experiments with natural groups (e.g. Goette et al., 2006, 2012) and groups formed in the laboratory (e.g. Abbink et al., 2012) demonstrate how competition between teams can help to overcome the downward trend in cooperation that is typically observed in public goods games. Yet, not only competition but also the mere provision of comparison information about an independent other group can foster within-group cooperation (Böhm et al., 2013; Burton-Chellew and West, 2012; Cárdenas and Mantilla, 2015; Sausgruber, 2009; Tan and Bolle, 2007). The cooperation-enhancing effect of between-group comparisons can be attributed to increased salience of group identity, resulting from a desire for positive distinctiveness and self-view (Mazar et al., 2008; Tajfel and Turner, 1979).[1]

Previous research has focused on (the benefits of) comparisons between symmetric groups. However, natural groups differ vastly on various domains. In many cases, individuals in different groups face different returns from cooperation. For instance, teams differ in terms of task complexity and team heterogeneity (Jehn et al., 1999). Additionally, an equivalent amount of cooperation effort may yield greater returns in a team working on tasks with a high stake than in a team working on routine tasks. Often, group members will know about the different returns from cooperation in other groups in their environment and, nevertheless, compare their

[1] Even in a setting with symmetric between-group comparisons, contributions to the public good may deteriorate if a comparison group provides a bad example, for instance by contributing very little.

own group's performance to that of other groups. Importantly, most social comparisons occur automatically, also to others who are not directly comparable (Gilbert et al., 1995; Verduyn et al., 2015). Apart from comparisons that occur automatically, a manager might consider making comparison information salient in order to increase team performance in teams with different returns from cooperation.

In this paper, we thus ask how information about different returns from cooperation in different groups affects individual decision processes when between-group comparison information is available. We thereby provide the first empirical test of how individuals' perception of differences in between-group comparisons affects within-group cooperation behavior. Psychological research on the perception of social groups (Ames, 2004; Corcoran and Mussweiler, 2009) suggests that social comparisons between structurally different groups might backfire by tainting the own group's perception, e.g., by increasing the salience of the own group's negative characteristics vis-à-vis the comparison group.

We aim at delineating the conditions under which – intended or unintended – comparisons between structurally different groups could have detrimental effects on within-group performance. To this end, we experimentally investigate a public goods game with comparison information about another group with a different return from cooperation, i.e., a different marginal per capita return. To answer the question of how the perception of differences between groups affects within-group cooperation, we vary whether or not groups have information about the structural difference between their own group and their comparison group. We suspect that only when group members have been made aware of the differences between groups, might this lead them to perceive between-group comparisons to be inadequate, and, in turn, might affect within-group cooperation.

Our findings support the view that, when individuals have information about the differences between groups, within-group cooperation becomes more fragile. We still observe that the desire to avoid 'lagging behind' the comparison group motivates contributions. Yet, we also find important limitations to between-group comparisons: Having information about structural between-group differences leads members of disadvantaged groups (i.e., groups with lower returns from cooperation) to reduce their contributions more strongly in response to unfavorable within-group comparisons. An analysis of inter-individual heterogeneity reveals that in particular those individuals who agree that the comparison group is not comparable to the own group decrease their contributions more strongly in response to unfavorable within-group comparisons.

In the following, we first describe our experimental design and derive our hypotheses regarding comparison processes among structurally different groups. Finally, we present our results in more detail, which are discussed afterwards.

2.2 Materials and Methods

2.2.1 Experimental Design

We study the effect of between-group comparisons when groups are structurally different in a public goods game. Two groups, g and h, with four players each, both play a repeated linear public goods game for 20 rounds (using a partner matching protocol). The groups play independently, i.e., one group cannot directly influence the other group's payoff. We assign either blue or green identities to groups.[2] In each round t, all players of both groups receive an identical endowment of $e = 20$. Each player i of group g decides on a contribution $c_{i,g,t} \in [0, e]$ to her group account. Aggregate group contributions per round $\sum c_{j,g,t}$ are redistributed equally to every group member and multiplied by the marginal per capita return (MPCR) μ_g. $\mu_g < 1$ constitutes free-riding incentives of group members. Player i's payoff in round t is $\pi_{i,g,t} = e - c_{i,g,t} + \mu_g \sum c_{j,g,t}$. The same rules apply to players of group h. Players of group g receive information on the average contributions of their comparison group h, $\bar{c}_{h,t-1}$, and their own group g, $\bar{c}_{g,t-1}$, in the preceding round $t - 1$, and vice versa.

To implement structural differences, one group faces a high MPCR ($\mu_g = 0.8$), while the other group has a low MPCR ($\mu_h = 0.4$). Empirically, a higher MPCR facilitates higher cooperation levels although free-riding incentives remain unchanged (Isaac and Walker, 1988). As additional controls, we have two further treatments where both groups have the same MPCR ($\mu_g = \mu_h = 0.4$).

To isolate how the perception of differences between groups affects within-group cooperation behavior, we additionally vary whether players only know the MPCR of their own group (NOINFO) or the MPCRs of both groups (INFO). In INFO treatments, the instructions explicitly informed players about the MPCR of the own group and the MPCR of the comparison group. By contrast, instructions in NOINFO treatments only contained information about the MPCR of the own group but did not contain any information about the MPCR of the comparison group. These variations comprise six experimental treatments,

[2] We rely on minimal, artificial groups to exclude group-specific stereotypes and a history of conflict (Tajfel et al., 1971). However, meta-analytical evidence suggests that the average effect size of ingroup favoritism does not differ in interactions between members from natural vs. artificial groups (Balliet et al., 2014).

summarized in Table 2.1.

Structural difference between groups	Is the MPCR of comparison group known?	
	yes	no
Own group MPCR $\mu_g = 0.4$ Comparison group MPCR $\mu_h = 0.8$	LOW-INFO	LOW-NOINFO
Own group: MPCR $\mu_g = 0.8$ Comparison group: MPCR $\mu_h = 0.4$	HIGH-INFO	HIGH-NOINFO
Both groups: MPCR $\mu_g = \mu_h = 0.4$	SAME-INFO	SAME-NOINFO

Table 2.1: Experimental treatments.

2.2.2 Participants and Procedure

We recruited N = 384 students via ORSEE (Greiner, 2015), resulting in 96 groups with four players each. The experiment was conducted in zTree (Fischbacher, 2007). Participants took part in experimental sessions of 32 players each, with each session being randomly assigned to one of the experimental conditions. There were two separate rooms, and players in different rooms were connected via the software. Participants of the same group were seated in the same room; participants of the respective comparison groups were seated in the other room. Because there were four groups in each room, participants could infer neither the identity of the other players in their own group nor the identity of players in their matched comparison group.

Participants received written instructions for the public goods game (see Appendix 2.A). Additionally, instructions were read aloud to ensure common knowledge. Participants had to answer questions of understanding before they actually started playing. In the INFO treatments, participants' questions also related to the comparison group's MPCR (same MPCRs in SAME-INFO and different MPCRs in LOW-INFO/HIGH-INFO, respectively). In NOINFO, questions related only to the MPCR of the own group.

After the 20 game-playing rounds, participants completed a questionnaire assessing their perceptions regarding the interaction situation on 7-point-scales (for details, see Appendix 2.B). We asked participants about their motivations for contributing to their group account. Further, we elicited the degree to which participants identify with their own group by using three items (Cronbach's $\alpha = 0.82$, Doosje et al. (1995), e.g., "I identify as a member of my group."). To assess whether participants consider the comparison group to be sufficiently similar for comparing

contributions between groups, we elicited their agreement with the statement: "The contributions of the other group are not comparable to the contributions in my group." We also assessed whether participants experience joy (or anger) when their group lies ahead (or lags behind) the comparison group.

After completing the post-experimental questionnaire, participants received their final payoff information and were paid in private. Participants earned on average about 18 Euro in sessions that lasted about 60 - 90 minutes.

2.3 Hypotheses

Social psychological research suggests that comparisons between structurally different groups may enhance the desire for positive distinctiveness of the own group, i.e., the status of the own group vis-à-vis the comparison group (Tajfel and Turner, 1979; Turner, 1975). When groups are symmetric and between-group information is available, research on public goods games has confirmed that individuals desire not to 'lag behind' a comparison group (Böhm et al., 2013; Burton-Chellew and West, 2012; Cárdenas and Mantilla, 2015; Sausgruber, 2009; Tan and Bolle, 2007). This desire, in turn, increases individuals' willingness to contribute to the provision of their group's public good. The desire to obtain positive distinctiveness is thus a potent motivational force and we expect it to operate regardless of whether individuals have or do not have information about their structural disadvantages, i.e., their lower return from cooperation relative to the comparison group. More precisely, individuals in both LOW-INFO and LOW-NOINFO should increase their contributions when their own group has contributed less than the comparison group, in an attempt to restore positive distinctiveness. Conversely, the desire for positive distinctiveness may be reduced when the group has contributed more than the comparison group, so in this case individuals should decrease their contributions. In the following, we focus on groups who lag behind their comparison group, because, from a group identity perspective, lying 'ahead' might also boost positive distinctiveness, whereas 'lagging behind' is an unambiguous threat to positive distinctiveness.

Hypothesis 1 (LOW): *Individuals in structurally disadvantaged groups increase their contributions when lagging behind the comparison group in terms of contributions.*

For individuals in structurally advantaged groups, i.e., individuals in HIGH-INFO and HIGH-NOINFO, the comparison group may be a less relevant comparison standard. This is because the own group typically lies ahead of the comparison group, so individuals may already feel distinct

in a positive way. Nonetheless, we expect individuals to be motivated by their desire not to lag behind the comparison group (albeit to a lesser extent).

Hypothesis 1 (HIGH): *Individuals in structurally advantaged groups increase their contributions when lagging behind the comparison group in terms of contributions.*

When a comparison group is perceived as different from the own group, this may not only affect between-group comparisons but may also taint the individuals' perception of their own group. According to the theory of 'optimal distinctiveness' (Brewer, 1991; Brewer and Weber, 1994), people's need for positive distinctiveness crucially determines how they attend to social information. Hence, when comparison groups are perceived as sufficiently structurally similar, between-group comparison processes should become important. In this case, individuals may expect to obtain positive distinctiveness by boosting their group identity. By contrast, when the own group is perceived as structurally disadvantaged, comparison processes between individuals of the same group should become more important. In LOW-INFO, individuals most likely expect their own group to perform worse in terms of contributions. As they dislike lagging behind the comparison group, they may focus more strongly on themselves compared to others within their group in order to achieve positive distinctiveness, i.e., they focus on personal identity rather than on group identity. It follows that within-group comparisons should become more salient.[3,4]

In within-group comparisons, conditional cooperation has been shown to play a crucial role (Fischbacher et al., 2001; Kelley and Stahelski, 1970). Conditional cooperation is often selfishly biased, meaning that unfavorable within-group comparisons receive more weight than favorable within-group comparisons (Fischbacher and Gächter, 2010). Thus, we expect individuals who have information about their group's structural disadvantage to respond more harshly particularly to unfavorable within-group comparisons.

Hypothesis 2 (LOW): *If individuals have information about the structural disadvantage of their group, they reduce their contribution more strongly in response to unfavorable within-group comparisons than if they do not have information about their disadvantage.*

[3]This does not mean that between-group comparisons should become less self-relevant. By contrast, a precondition for a stronger focus on personal identity is that people who receive information about disadvantages in their group find the group comparison to be self-relevant and dislike lagging behind their comparison group.

[4]An alternative explanation for the increased salience of within-group comparisons may relate to the selective accessibility model of social comparisons (Mussweiler, 2003). This model predicts that after an initial assessment of similarity, individuals engage in a selective search, confirming either similarities or differences: Individuals who initially expect chances of cooperation to be lower in the own group may attend more strongly to unfavorable within-group comparisons (to confirm their initial expectation of low cooperation). Note that while within-group comparisons should become more salient, between-group comparisons should not, because information about a group's disadvantage also makes it easier to explain why the own group lags behind the comparison group.

For structurally advantaged groups, information about the structural advantage reflects positively on their group's distinctiveness. Consequently, contrary to disadvantaged groups, we do not expect individuals to focus on personal identity when receiving information about their advantages.[5] Yet, the structural advantage also makes it more difficult to attribute relative successes to the own group's positive distinctiveness alone rather than to external circumstances. Hence, we expect having information about the own group's structural advantage to neither positively nor negatively affect reactions to within-group comparisons.

Hypothesis 2 (HIGH): *If individuals have information about the structural advantage of their group, they do not reduce their contribution more strongly in response to unfavorable within-group comparisons than if they do not have information about their advantage.*

In our design, the key manipulation is to provide (or not to provide) information about the dimension of similarity, and we expect perceptions of similarity to be affected regardless of whether individuals have information about structural advantages or structural disadvantages. A wide array of psychological research shows that perceived similarity is one of the most important moderators of social comparisons (Ames, 2004; Corcoran et al., 2011). At the same time, there is also considerable evidence suggesting that people use social information in a self-serving way, especially when the information is highly self-relevant (Kunda, 1990; Pyszczynski et al., 1985; Taylor et al., 1996). Hence, particularly if individuals have information about their structural disadvantage – which may potentially threaten positive distinctiveness – they may use the information in a self-serving way to argue that groups are not sufficiently similar for contributions to be comparable between groups. Members of disadvantaged groups may thus feel that contributions are not comparable (and perhaps ought not to be compared) between groups, and thus they focus more strongly on how they compare to others in their team, i.e., their personal identity.[6,7] We thus hypothesize that the individual's (ex-post) perception that contributions are not comparable between groups should play a key role with regard to how strongly individuals reduce their contributions in response

[5] This is also in accordance with models of inequity aversion, where disadvantageous inequities are strictly more aversive than advantageous inequities (Fehr and Schmidt, 1999).

[6] This is not to say that individuals who perceive that groups are not comparable also attend less to between-group comparison information or that their desire to keep up with the comparison group is hampered (this would contradict Hypothesis 1). Instead, the argument here is that individuals use information about differences between groups to maintain a positive self-view when focusing on their personal identity.

[7] Alternatively, individuals who perceive that contributions are not comparable between groups may simply focus on within-group comparisons as a result of their selective search for differences rather than for similarities. In this case, we need to assume that the desire for positive group distinctiveness operates regardless of perceived similarity, so between-group comparisons are not affected by an individual's perception that contributions are not comparable between groups.

to unfavorable within-group comparisons:

Hypothesis 3 (LOW): *If they have information about the structural advantage of their group, individuals who agree more strongly that contributions are not comparable between groups should reduce their contributions more strongly in response to unfavorable within-group comparisons. This relation should be stronger in groups that have information about their disadvantage than in groups that do not have information about their disadvantage.*

Self-serving information processing should be present in disadvantaged but not in advantaged groups because only information about structural disadvantages threatens the own chances of achieving positive (group) distinctiveness. For a structurally advantaged group, between-group comparisons are less self-relevant and less threatening because such a group is continuously ahead of its comparison group. For this reason, we do not expect an individual's perception that contributions are not comparable between groups to predict that an individual will focus on personal identity.

Hypothesis 3 (HIGH): *If they have information about the structural advantage of their own group, individuals who agree more strongly that contributions are not comparable between groups should not reduce their contributions more strongly in response to unfavorable within-group comparisons. This relation should be similar in groups that do not have information about their advantage.*

See Table 2.2 for an overview of our hypotheses.

	Hypotheses
H1 (LOW)	In disadvantaged groups: Contributions increase when the own group lags behind the comparison group.
H1 (HIGH)	In advantaged groups: Contributions increase when the own group lags behind the comparison group.
H2 (LOW)	Contribution reductions in response to unfavorable within-group comparisons: LOW-INFO $>$ LOW-NOINFO
H2 (HIGH)	Contribution reductions in response to unfavorable within-group comparisons: HIGH-INFO $\approx$ HIGH-NOINFO
H3 (LOW)	Contribution reductions in response to unfavorable within-group comparisons are stronger for individuals who agree more strongly that contributions are not comparable between groups: LOW-INFO $>$ LOW-NOINFO
H3 (HIGH)	Contribution reductions in response to unfavorable within-group comparisons are not stronger for individuals who agree more strongly that contributions are not comparable between groups: HIGH-INFO $\approx$ HIGH-NOINFO

Table 2.2: Overview of Hypotheses.

2.4 Results

On the aggregate level, groups with a low return from cooperation (MPCR $\mu = 0.4$, LOW-INFO and LOW-NOINFO) show a similar downward trend in cooperation over rounds, whereas groups with a higher return from cooperation (MPCR $\mu = 0.8$, HIGH-INFO and HIGH-NOINFO) show only a modest downward trend in contributions, so contributions stay at a much higher level (see also Figure 2.1 and Appendix 2.C for descriptive statistics). This observation is in line with previous research showing that a higher MPCR leads to higher contributions, even though free-riding incentives remain unchanged (Isaac and Walker, 1988).

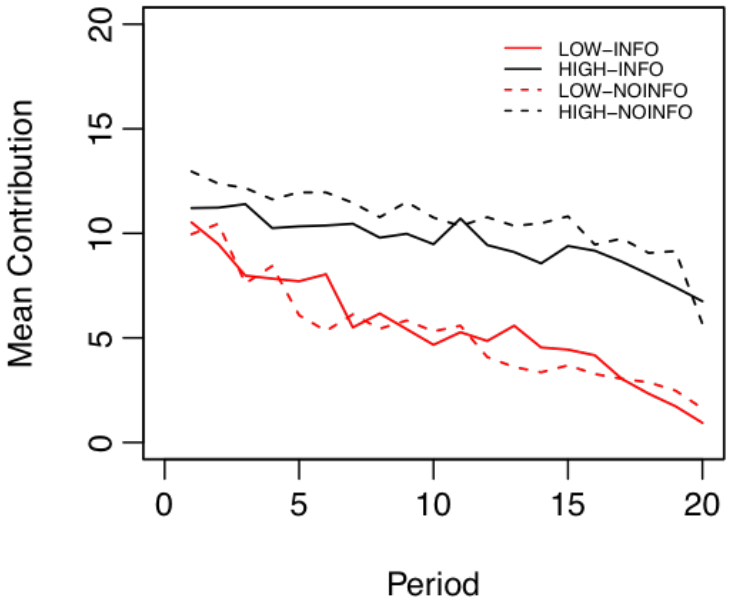


Figure 2.1: Group-aggregated contribution levels across rounds.

To see how information about between-group differences in returns from cooperation affects the individual underlying decision processes, we first look at group-aggregates to foreshadow our main results. Namely, individuals who have information about their group's structural disadvantage have the tendency to reduce their contributions more strongly when members of the own group contribute less than the individual (unfavorable within-group comparisons) compared to individuals who do not have information about their group's disadvantage (MWU: LOW-INFO vs. LOW-NOINFO, mean -3.74 vs. -2.83, $Z = 1.85$, $p = .068$; cf. Result 2 (LOW)). If we further restrict analyses to individuals who agree at median-levels or above within their respective treatment that contributions are not comparable between groups, individuals who have information about their group's disadvantage show significantly stronger responses to unfavorable comparisons compared to individuals who do not have information about their

group's disadvantage (MWU: LOW-INFO vs. LOW-NOINFO, mean -5.22 vs. -3.34, $Z = 2.62$, $p = .008$, cf. Result 3 (LOW)).[8] This is first evidence that having information about structural disadvantages makes individuals less tolerant of unfavorable within-group comparisons, i.e., of within-group free-riding. Further, similarity perceptions seem to play a key role in explaining stronger responses to within-group free-riding.

To investigate formally whether having information about structural differences in between-group comparisons, i.e., information about having different MPCRs, affects individual contribution adaptations in repeated within-group cooperation behavior, we estimate several mixed-effect models. Our models include random effects to account for dependencies between paired groups and within groups (due to feedback information after each round), and at the individual level (due to repeated decisions over rounds). We refer the reader to Appendix 2.D for the equations of our statistical models. As a dependent variable, we analyze the players' contribution change from the actual to the previous round (*Contribution Change*). We analyze behavior for players with MPCR $\mu = 0.4$ and MPCR $\mu = 0.8$ separately. Thus, players who are statistically compared face equal incentives across treatments. In Table 2.3, we analyze behavior in groups with MPCR $\mu = 0.4$ by presenting Models L0-L5 (the prefix "L" indicates MPCR $\mu = 0.4$). In Table 2.5 in the Appendix, we additionally analyze behavior in groups with MPCR $\mu = 0.8$ by presenting Models H0-H4 (the prefix "H" indicates MPCR $\mu = 0.8$) analogously to Models L0-L4.

Our models account for behavioral patterns known from the literature by estimating different fixed effects. As a measure of conditional cooperation (e.g., Fischbacher et al., 2001), we use a dummy variable indicating whether average contributions by group members were lower than an individual's contribution in the preceding round (*Unfavorable Comparison*).[9] To measure the impact of between-group comparisons (e.g., Böhm et al., 2013), we use a dummy variable indicating whether contributions in the preceding round in the own group were lower than in the comparison group (*Group Behind Comparison Group*).[10] Moreover, our models include

[8] Re-estimating this test for individuals who agree at below median-levels in their respective treatment that contributions are not comparable between groups yields no significant differences to individuals in disadvantaged groups who do not have information about their disadvantages (MWU: LOW-INFO vs. LOW-NOINFO, mean -3.26 vs. -2.99, $Z = 0.68$, $p = .529$).

[9] Please note that individuals who contribute zero to the public good can avoid receiving an unfavorable within-group comparison. This raises a selection issue which might affect our coefficient estimates for *Unfavorable Comparison*. If we exclude individuals who contribute zero from our analysis, however, our coefficient estimates for Unfavorable Comparison are barely affected.

[10] We do not include a variable *Group Ahead Comparison Group*, since this variable is almost perfectly collinear with Group Behind Comparison Group (Spearman's $\rho = -.94$). There are only 3 out of 240 periods in LOW-INFO groups and 3 out of 240 in LOW-NOINFO groups where average contributions are the same as for the group with higher MPCR (SAME-INFO: 16/240, SAME-NOINFO: 5/240).

treatment dummies to address unexplained heterogeneity between treatments (with *LOW-INFO* as the reference category in Models L0-L5; with *HIGH-INFO* as the reference category in Models H0-H4). We also include *Round* as a time trend indicator.

In Models L0 and H0, we include only treatment dummies and the variable *Group Behind Comparison Group* to estimate the 'full' effect of between-group comparisons on contributions. In contrast, Models L1 and H1 also control for conditional cooperation and a time trend. Consistent with studies investigating between-group comparisons in public goods games, we find evidence for the motivating force of positive group distinctiveness, especially in structurally disadvantaged groups (Hypothesis 1 (LOW)).

Result 1 (LOW): *If the own group is structurally disadvantaged, individuals increase their contributions when lagging behind the comparison group ($\beta_{\text{Group Behind Comparison Group}}$, Model L0, $p < .001$, Models L1-L2, p-values $< .10$, Models L3-L4, p-values $< .05$, Model L5, $p < .001$).*

Result 1 (HIGH): *If the own group is structurally advantaged, individuals do not systematically increase their contributions when lagging behind the comparison group ($\beta_{\text{Group Behind Comparison Group}}$, Model H0, $p < .10$, Models H1-H4, p-values $> .10$).*[11]

In Models L2 and L3, we analyze how the information about the structural disadvantage, i.e., about having a lower MPCR, changes players' reactions to within-group behavior. We extend Model L1 by adding interactions between *Unfavorable Comparison* and treatment dummies. As before, we use LOW-INFO as a reference group, so the coefficient $\beta_{\text{Unfavorable Comparison}}$ (without treatment interactions) shows how individuals in LOW-INFO change their contributions in response to unfavorable within-group comparisons. The treatment interactions, e.g., $\beta_{\text{Unfavorable Comparison*LOW-NOINFO}}$, further reveal how responses to unfavorable within-group comparisons in LOW-INFO differ compared to other treatments, e.g., LOW-NOINFO. Positively signed interaction coefficients indicate that reductions in response to unfavorable within-group comparisons are stronger in LOW-INFO compared to the other treatments. In Model L3, we additionally include a variable indicating whether other group members contributed more on average than the individual in the preceding round (*Favorable Comparison*). We include this variable because it can be interpreted as measuring positive

[11] We find partial support for Hypothesis 1 (HIGH), when we extend our models by the variable Comparison Group Minus Group measuring the continuous difference between the comparison groups' and the own group's average contribution (see Table 2.7 and 2.8 in the Appendix, Models L1'-L4' and H1'-H4'). In particular, we observe that for individuals in advantaged groups, contributions increase as between-group differences increase ($\beta_{\text{Comparison Group Minus Group}}$, Models H1'- H4', all p-values $< .05$).

reciprocity, i.e., responses from reciprocating free-riders, while *Unfavorable Comparison* can be interpreted as measuring negative reciprocity, i.e., responses to others' free-riding behavior.[12] We conduct analagous analyses for structurally advantaged players in Models H2 and H3.

Consistent with our central hypothesis (Hypothesis 2 LOW), individuals' sensitivity to unfavorable within-group comparisons increases if they have information about their own group's disadvantage, but not if they have information about their own group's advantage (Hypothesis 2 HIGH):

Result 2 (LOW): *Individuals who have information about their group's structural disadvantage (LOW-INFO) reduce their contributions considerably more strongly in response to unfavorable within-group comparisons (by 39%, Model L3) than individuals who do not have information about their disadvantage (LOW-NOINFO) ($\beta_{Unfavorable\ Comparison*LOW\text{-}NOINFO}$, Models L2-L3, p-values $< .01$).*

Result 2 (HIGH): *Individuals who have information about their group's structural advantage (HIGH-INFO) do not reduce their contributions significantly more strongly than individuals who do not have information about their advantage (HIGH-NOINFO) ($\beta_{Unfavorable\ Comparison*LOW\text{-}NOINFO}$, Models H2-H4, p-values $> .10$).*

This suggests that particularly if individuals have information about their structural disadvantage, they become less tolerant of within-group free-riding, so that within-group cooperation becomes more fragile. Figure 2.2 plots the cumulative distribution of individuals' average contribution reductions in response to unfavorable within-group comparisons separately for the treatments LOW-INFO and LOW-NOINFO. The figure supports Result 2 (LOW) by showing that harsh reactions to unfavorable within-group comparisons are more frequent in LOW-INFO than in LOW-NOINFO. Importantly, contribution reductions in response to unfavorable within-group comparisons are also stronger in LOW-INFO than in treatments without structural differences, regardless of whether individuals have information (LOW-INFO vs. SAME-INFO: $\beta_{\text{Unfavorable Comparison*SAME-INFO}}$, Models L2-L3, p-values $< .05$, reduction difference 25% in Model L3) or do not have information about facing the same MPCR (LOW-INFO vs. SAME-NOINFO: $\beta_{\text{Unfavorable Comparison*SAME-NOINFO}}$, Models L2-L3, p-values $< .01$, reduction difference 33% in Model L3).

[12] Individuals contributed the same as the group average in 15% of the cases in LOW-INFO (LOW-NOINFO: 11%, SAME-INFO: 19%, SAME-NOINFO: 17%; HIGH-INFO: 9%; HIGH-NOINFO: 15%), so collinearity between Favorable Comparison and Unfavorable Comparison is unproblematic (Spearman's ρ = -.72). Favorable Comparison significantly increases model fit, so we retain it in Models L4-L5. Dropping it does not change significance patterns in ways affecting our results.

	Contribution Change					
	(L0)	(L1)	(L2)	(L3)	(L4)	(L5)
LOW-NOINFO	0.004	0.120	−0.320	−0.377	−0.493	−1.038
	(0.225)	(0.218)	(0.274)	(0.262)	(0.318)	(0.700)
SAME-INFO	0.210	0.032	−0.305	−0.242	−0.322	−0.811
	(0.197)	(0.192)	(0.237)	(0.227)	(0.283)	(0.633)
SAME-NOINFO	0.221	0.124	−0.290	−0.238	−0.298	−0.982
	(0.198)	(0.191)	(0.238)	(0.228)	(0.284)	(0.632)
Group behind Comparison Group	0.736****	0.217*	0.223*	0.265**	0.278**	0.565****
	(0.136)	(0.129)	(0.130)	(0.127)	(0.127)	(0.162)
Round		−0.038****	−0.038****	−0.023**	−0.024**	−0.030***
		(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Unfavorable Comparison		−4.103****	−4.970****	−4.224****	−3.908****	−4.263****
		(0.125)	(0.302)	(0.331)	(0.385)	(0.914)
Unfavorable Comparison*LOW-NOINFO			1.154***	1.195***	1.004**	0.872
			(0.424)	(0.422)	(0.484)	(1.045)
Unfavorable Comparison*SAME-INFO			0.903**	0.855**	0.620	0.443
			(0.370)	(0.368)	(0.419)	(0.934)
Unfavorable Comparison*SAME-NOINFO			1.101***	1.052***	0.719*	0.995
			(0.369)	(0.367)	(0.418)	(0.935)
Favorable Comparison				0.906****	0.893****	0.892****
				(0.176)	(0.176)	(0.182)
Not Comparable					−0.037	0.015
					(0.108)	(0.127)
Not Comparable*LOW-NOINFO					−0.049	−0.082
					(0.135)	(0.162)
Not Comparable*SAME-INFO					−0.047	−0.139
					(0.125)	(0.145)
Not Comparable*SAME-NOINFO					0.020	−0.029
					(0.127)	(0.150)
Unfavorable Comparison*Not Comparable					−0.412***	−0.743****
					(0.157)	(0.199)
Unfavorable Comparison*Not Comparable* LOW-NOINFO					0.570***	0.780***
					(0.210)	(0.266)
Unfavorable Comparison*Not Comparable* SAME-INFO					0.693****	1.111****
					(0.189)	(0.230)
Unfavorable Comparison*Not Comparable* SAME-NOINFO					0.398**	0.728***
					(0.190)	(0.238)
Interactions with other questionnaire variables are omitted for brevity in this table.						
Constant	-1.031****	1.307****	1.629****	0.727**	0.792**	1.452**
	(0.187)	(0.228)	(0.251)	(0.301)	(0.344)	(0.663)
Observations	5,472	5,472	5,472	5,472	5,472	5,472
Log Likelihood	−16,356	−15,870	−15,865	−15,852	−15,840	−15,902
Model Fit Increase (L-Ratio: X^2)	-	970.7****	10.4**	26.1***	24.4***	176.6****

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; **** $p < 0.001$

Table 2.3: Mixed-effects models predicting contribution change. Clustering is at the level of a pair-of-groups, the group, and the individual. LOW-INFO is the reference category. In Models L1-L5, the model fit does not significantly increase by adding interactions between treatment dummies and Favorable Comparison (e.g., L-Ratio test against Model L3, $X^2(3) = 0.26$, $p = 0.964$) or Group Behind Comparison Group (e.g., L-Ratio test against Model L3, $X^2(3) = 1.22$, p = 0.748). See Appendix Table 2.6, for the corresponding Models L3a and L3b.

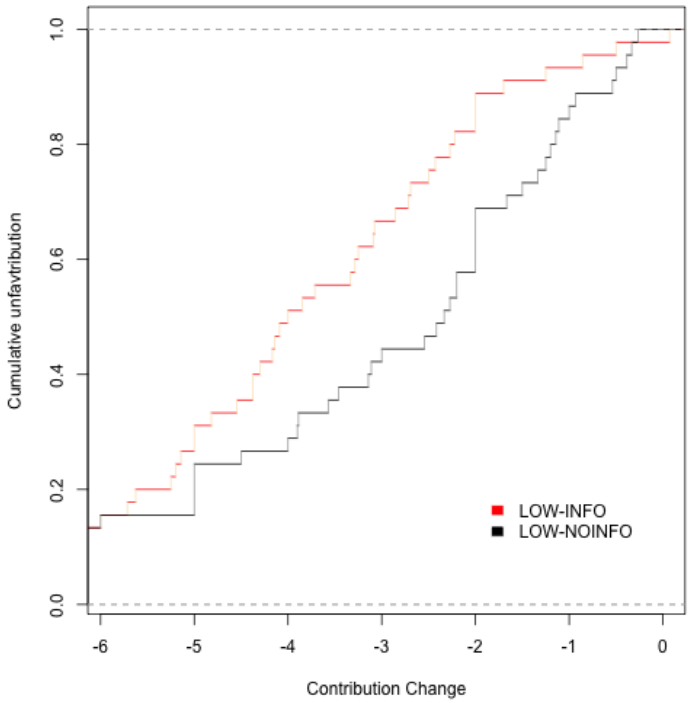


Figure 2.2: Cumulative distribution of individual contribution change in response to unfavorable within-group comparisons.

In Models L4 and L5, we investigate whether information about structural disadvantages has a stronger effect on participants who agree that contributions are not comparable between groups (*Not Comparable*). We extend Model L3 by interacting *Not Comparable* with *Unfavorable Comparison* and treatment dummies. As before, we use LOW-INFO as a reference group, so the coefficient $\beta_{\text{Unfavorable Comparison*Not Comparable}}$ reveals – for the treatment LOW-INFO – the relation between responses to unfavorable within-group comparisons and an individual's agreement that contributions are not comparable between groups. The treatment interactions, e.g., $\beta_{\text{Unfavorable Comparison*Not Comparable*LOW-NOINFO}}$, further reveal how this relation differs from LOW-INFO in other treatments. We repeat analogous analyses in Model H4 for advantaged groups. We find support for Hypothesis 3 (LOW):

Result 3 (LOW): *Individuals who agree more strongly that contributions are not comparable between groups also reduce their contributions more strongly (in LOW-INFO,* $\beta_{\textit{Unfavorable Comparison*Not Comparable}}$*, Models L4-L5, p-values* $< .001$*) and they reduce their contributions more strongly if they have information about their structural disadvantage than if they do not have this information (LOW-NOINFO,* $\beta_{\textit{Unfavorable Comparison*Not Comparable*LOW-NOINFO}}$*, Models L4-L5,* $p < .01$*).*

Contrary to disadvantaged groups, we find that the perception that contributions are not comparable between groups does not predict reductions in response to unfavorable within-group comparisons in advantaged groups, which supports Hypothesis 3 (HIGH):

Result 3 (HIGH): *Individuals in structurally advantaged groups who agree more strongly that contributions are not comparable between groups do not reduce their contributions more strongly in response to unfavorable within-group comparisons, regardless of whether they have information or do not have information about their structural advantage (e.g.,* $\beta_{Unfavorable\ Comparison*Not\ Comparable}$*, Model H4,* $p > .50$*).*

Figure 2.3 corroborates Result 3 (LOW) by showing that individuals with above-median (modal) agreement (*Not Comparable*: Median = 7) display much harsher reactions to unfavorable within-group comparisons than those below the median. Moreover, agreeing that contributions are not comparable between groups also affects contribution reductions more strongly in LOW-INFO than in treatments without structural differences, regardless of whether individuals have information (SAME-INFO, $\beta_{\text{Unfavorable Comparison*Not Comparable*SAME-INFO}}$, Model L4-L5, $p < .05$) or do not have information about facing the same MPCR (SAME-NOINFO, $\beta_{\text{Unfavorable Comparison*Not Comparable*SAME-NOINFO}}$, Model L4-L5, $p < .001$). Figure 2.4 in the Appendix shows the distribution of the variable *Not Comparable* for groups with a lower MPCR (LOW-INFO, LOW-NOINFO, SAME-INFO, SAME-NOINFO) and groups with a higher MPCR (HIGH-INFO and HIGH-NOINFO). Interestingly, in LOW-INFO the mode of the variable *Not Comparable* is 7, whereas in LOW-NOINFO the mode of the variable *Not Comparable* is 1. For HIGH-INFO and HIGH-NOINFO groups, there is a similar but less pronounced shift in distributions. This suggests that the provision of information about structural differences greatly changes players' perceptions that contributions are not comparable between groups. For an overview of descriptive questionnaire data for each treatment, we refer the reader to Appendix 2.G.

In Model L5, we account for the multiple measures in our questionnaire, which makes it more likely for differences to be detected by chance (multiple comparisons). We interact each of the questionnaire measures with *Unfavorable Comparison* and treatment dummies. For each questionnaire measure, we can thus infer how responses to unfavorable within-group comparisons depend on the questionnaire measure and how this relation differs from LOW-INFO in other treatments. To obtain adjusted p-values for the 42 possible comparisons (14 questionnaire measures x 3 control treatments), we apply a step-down Dunnett procedure.

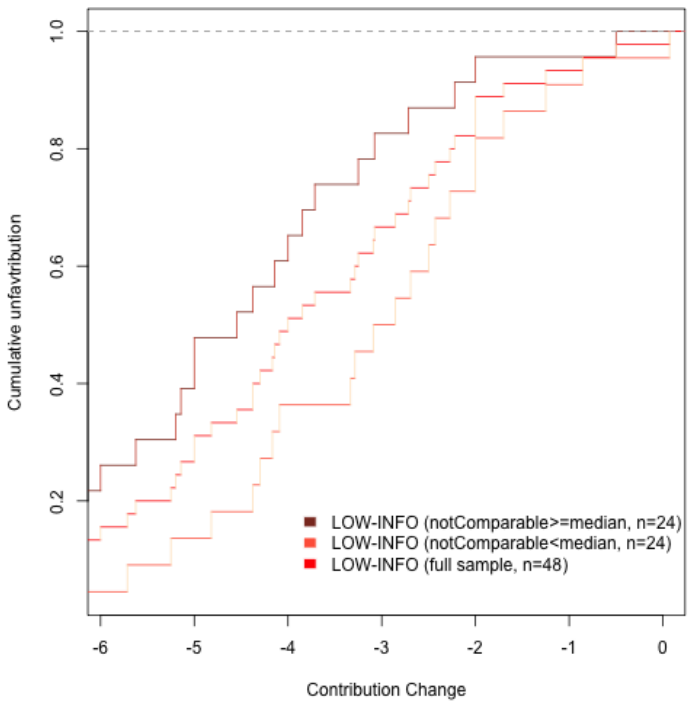


Figure 2.3: Cumulative distribution of individuals' contribution change in response to unfavorable within-group comparisons conditional on whether individuals agree that contributions of groups are not comparable between groups (*Not Comparable*).

The finding from Result 3 (LOW) (i.e., that an individual's agreement that contributions are not comparable between groups explains why this individual reacts more harshly to unfavorable comparisons in LOW-INFO but not in LOW-NOINFO) is robust to this very conservative procedure at a 10% level (LOW-NOINFO, $\beta_{\text{UnfavorableComparison*Not Comparable*LOW-NOINFO}}$, adjusted p = .092, SAME-NOINFO, $\beta_{\text{UnfavorableComparison*Not Comparable*SAME-INFO}}$, adjusted p = .062, SAME-INFO, $\beta_{\text{UnfavorableComparison*Not Comparable*SAME-NOINFO}}$, adjusted $p < .001$).

While there are no robust interaction effects for other questionnaire measures, their assessment still yields interesting insights. By and large, emotional reactions and perceptions in the decision situation affect contribution adaptations in expected directions: Individuals experiencing anger (joy) over their group's poor (good) relative performance decrease (increase) their contributions more strongly (β_{Anger} = -0.300, β_{Joy} = 0.318, both p-values $< .05$). Moreover, individuals who perceive the decision situation as a between-group comparison show harsher contributions reductions in response to unfavorable within-group comparisons ($\beta_{\text{Unfavorable Comparison*Between-Group Comparison}}$ = -0.306, $p < .10$). Perhaps most interestingly, individuals in LOW-INFO who report more joy when their group lies ahead reduce their contribution more strongly in response to unfavorable within-group comparisons (Model L5,

LOW-INFO as reference group, $\beta_{\text{Unfavorable Comparison*Joy}}$ = -0.466, $p < .10$) in comparison to individuals reporting more joy in LOW-NOINFO, SAME-INFO and SAME-NOINFO ($\beta_{\text{Unfavorable Comparison*Joy*LOW-NOINFO}} = 0.627$, $\beta_{\text{Unfavorable Comparison*Joy*LOW-NOINFO}} = 0.720$, $\beta_{\text{Unfavorable Comparison*Joy*LOW-NOINFO}} = 0.582$, all p-values $< .10$). This suggests that individuals in LOW-INFO, who find the between-group comparison to be more self-relevant, react more harshly to free-riding, potentially in order to motivate group members to 'stand together'.

2.5 Discussion

We investigate a pervasive and practically relevant condition of economic peer effects: comparisons between structurally different groups. Although we confirm the general pattern that the desire to avoid 'lagging behind' a comparison group motivates within-group cooperation in the public goods game, we report important limitations of between-group comparisons for fostering within-group cooperation. That is, if group members have information about the own group's structural disadvantage, they are more sensitive to how well their group 'stands together'. Compared to individuals who do not have information about their group's disadvantage, they reduce their contributions more strongly in response to unfavorable within-group comparisons.

Our data provide some hints on the underlying mechanism. In particular, the extent to which individuals agree that contributions are not comparable between groups predicts why some individuals react more harshly than others when faced with unfavorable within-group comparisons in disadvantaged groups. Notably, this finding is in line with the idea that individuals often use social comparison information in a self-serving way (Kunda, 1990; Pyszczynski et al., 1985; Taylor et al., 1996): For individuals in disadvantaged groups, the between-group comparison is highly self-relevant, because it potentially threatens the own group's positive distinctiveness, which is not the case for advantaged groups. In line with this, perceptions that contributions are not comparable between groups predict contribution reductions in response to unfavorable within-group comparisons for individuals who have information about their groups's disadvantages, but not for individuals who have information about their group's advantages. We also find some suggestive evidence that greater joy (when the own group lies ahead) predicts harsher contribution reductions in response to unfavorable comparisons, but again only for individuals who have information about their group's

disadvantages. Harsher contribution reductions in response to unfavorable comparisons might thus partly be responses from group members who desire to have the group 'standing together' and thereby increasing positive group distinctiveness.[13,14]

A limitation in identifying the psychological mechanism is that we rely on the individual's ex-post perception that contributions are not comparable between groups. This perception may be subject to ex-post rationalization after playing the game. However, we argue that both individuals who have information about their group's disadvantage and individuals who do not have this information experience their own group continuously making lower contributions than the comparison group. Consequently, individuals in both treatments could ex-post rationalize that the groups are different and are thus not comparable. Yet, since information about differences is provided before the game starts for individuals who have information, but not for those who do not have information, it is most likely that receiving information about differences affected individuals' initial perceptions of similarity and influenced their subsequent comparison processes.

An important caveat in interpreting our results relates to income differences associated with structural disadvantages: Providing information about lower returns from cooperation in one group relative to a comparison group necessarily implies that the former will earn less than the latter for a given level of contributions. From an inequity aversion perspective (Fehr and Schmidt, 1999), information about income differences stemming from different returns from cooperation should primarily affect how individuals perceive disadvantageous income differences stemming from between-group comparisons, but not how they perceive disadvantageous income differences stemming from within-group comparisons. This is however not supported by our data. Further, not only information about structural disadvantages but also information about structural advantages should make income differences between groups more aversive, which is also not supported by our data. From a psychological perspective, we think that it is nonetheless

[13]Interestingly, the selective accessibility model (Mussweiler, 2003) also predicts that perceived similarity crucially moderates how individuals repond to comparison information. In this model, individuals initially assess whether a comparison standard is similar or not and then they selectively search for information confirming either similarities or differences. Analogously, in our experiment, receiving information about differing returns from cooperation may induce a selective focus on what differentiates the own group from the other group, so individuals become more sensitive to (unfavorable) within-group comparisons. Conversely, advantaged groups should also respond to favorable comparisons more positively. We do not observe such a boost, so we prefer a motivational explanation here.

[14]An alternative explanation might relate to trust. Inequities due to differences in returns from cooperation may reduce trust and thus explain stronger contribution reductions (Anderson et al., 2008; Greiner et al., 2012). Our questionnaire measure that comes closest to 'trust' is a group's perceived "general willingness to contribute". Yet, contribution reductions in response to unfavorable comparisons do not differ depending on an individual's score on this measure between treatments (Model L5, $\beta_{\text{Unfavorable Comparison*OwnGroupCoop}}$, $p > .10$, $\beta_{\text{Unfavorable Comparison*OwnGroupCoop*LOW-NOINFO}}$, $p > .10$).

impossible to disentangle the effects of information about different returns from cooperation from the effects of information about income differences. It may well be that perceptions of income differences, not lower returns from cooperation, threaten an individual's feeling of positive group distinctiveness and thus induce a focus on personal identity. For our hypotheses, however, it does not matter whether perceptions of income differences or perceptions of different returns from cooperation drive the results. Instead, the decisive factor is whether or not being in a group with lower returns from cooperation threatens positive feelings of group distinctiveness.

Our findings offer a broad scope for extensions. Future research might assess whether comparisons to structurally different groups are unpleasant per se, e.g., by letting individuals choose between either a similar or a different comparison group. Further research could also explore individuals' motivations or willingness-to-pay for switching to a group which faces better conditions than the own group. More broadly, we believe that comparisons to different others have received too little attention in the study of economic peer effects. Yet, research in this direction could identify important nonlinearities by delineating, for example, when people contrast away from, rather than conform with, unethical peers.

2.6 Conclusion

If a comparison group is perceived to be different in terms of returns from cooperation, between-group comparisons may produce undesirable effects of rendering cooperation more fragile. This study highlights perceived similarity as a previously neglected psychological factor of economic peer effects between teams, which may be relevant for peer effects in many economic contexts.

Appendices

2.A Experimental instructions

Please note that instructions for the treatments LOW-INFO and LOW-NOINFO are included for illustration purposes. Instructions for other treatments and original, untranslated German instructions are available from the authors upon request.

Treatment LOW-INFO

General information

Assignment of groups

- Each participant in the experiment will be randomly assigned to either a **blue group** or a **green group**. Both groups will consist of **four participants** each.
- One blue group and one green group will be **randomly assigned to each other.**
- You will not know who belongs to your group or who belongs to the other group.
- The partitioning and the assignment of groups will remain unchanged over the course of the entire experiment. This means that the members of your group as well as the members of the other group which will be assigned to your group are the same persons during the entire experiment.

Rounds

- The experiment consists of **20 rounds** and the experiment is different for members of the blue and the green groups.
- Each round has the following stages:
  - At the beginning of each round, each participant receives an endowment of 20 points.
  - Each participant has the possibility to **contribute an integer amount of points between 0 and 20 to their own group project.** Points which a participant does not contribute to the group project will be kept by that participant.
  - **Each participant of the own group profits** from each contribution to the own group project.
  - **For members of the blue group:** Each point which is contributed to the project by members of the own group will be **multiplied by 3.2** and **redistributed equally among all participants of the own group. Accordingly, each of the four participants of the own group receives 0.8 points from each point, which has been contributed by any member of the own group.**

- **For members of the green group:** Each point which is contributed to the project by members of the own group will be **multiplied by 1.6** and **redistributed equally among all participants of the own group. Accordingly, each of the four participants of the own group receives 0.4 points from each point, which was contributed by any member of the own group.**
- Points which are **not contributed to the group project profit only the participant proper.**

The **earnings for members of blue groups** are calculated as follows:

Earnings for members of the blue group
= **kept points + 0.8** x sum of all **contributions of all members of the blue group** to the blue group‘s project

---

The **earnings for members of green groups** are calculated as follows:

Earnings for members of the green group
= **kept points + 0.4** x sum of all **contributions of all members of the green group** to the green group‘s project

---

Example:

Group	Blue				Green			
Member	B1	B2	B3	B4	G1	G2	G3	G4
Endowment	20	20	20	20	20	20	20	20
Contribution to the project	**0**	**20**	**16**	**8**	**16**	**8**	**4**	**20**
Sum of contributions	44				48			
Earnings from the project	0,8 x 44 = 35.2				0,4 x 48 = 19.2			
Kept	**20**	**0**	**4**	**12**	**4**	**12**	**16**	**0**
Earnings	**55.2**	**35.2**	**39.2**	**47.2**	**23.2**	**31.2**	**35.2**	**19.2**

Information about the contributions to the project

After completion of each contribution decision you receive the following information:

- Average contributions of members of the green group to the green group's project
- Average contributions of members of the blue group to the blue group's project
- Your earnings in the round

Payment

- Upon completion of the 20 decision rounds, you will be informed about your final earnings, i.e., the sum of the 20 per-round earnings, on your computer screen.
- All earned points will be converted and paid in cash according to the following exchange rate: **10 points = 0.25 Euro.**
- Afterwards, each participant will be called separately and earnings will be paid out in private and anonymously. You will not find out about the earnings of other participants; likewise, other participants will not find out about your earnings.

Treatment LOW-NOINFO

General information

Assignment of groups

- Each participant in the experiment will be randomly assigned to either a **blue group** or a **green group**. Both groups consist of **four participants** each.
- One blue group and one green group will be **randomly assigned to each other**.
- You will not know who belongs to your group or who belongs to the other group.
- The partitioning and the assignment of groups will remain unchanged over the course of the entire experiment. This means that the members of your group as well as the members of the other group which will be assigned to your group are the same persons during the entire experiment.

Rounds

- The experiment consists of **20 rounds**.
- Each round has the following stages:
  - At the beginning of each round, each participant receives an **endowment of 20 points**.
  - Each participant has the possibility **to contribute an integer amount of points between 0 and 20 to their own group project**. Points which a participant does not contribute to the group project will be kept by that participant.
  - **Each participant of the own group profits** from each contribution to the own group project.
  - **More precisely:** Each point, which is contributed to the project by members of the own group will be **multiplied by 1.6** and **redistributed equally among all participants of the own group. Accordingly, each of the four participants of the own group receives 0.4 points from each point, which has been contributed by any member of the own group.**
  - Points which are **not contributed to the group project profit only the participant proper.**

The **earnings** for members of **your group** are calculated as follows:

Earnings for members of your group
= **kept points + 0.4** x sum of all **contributions of all members of your group** to the group's project

---

Example:

Group	Green				Blue			
Member	G1	G2	G3	G4	B1	B2	B3	B4
Endowment	20	20	20	20	20	20	20	20
Contribution to the project	**0**	**20**	**16**	**8**	**16**	**8**	**4**	**20**
Sum of contributions	44				48			
Kept	**20**	**0**	**4**	**12**	**4**	**12**	**16**	**0**

Each member of the green group receives earnings from the green group's project which amount to 0.4 x 44 = 17.6 points.
Member G1 earns 20+17.6=37.6 points.
Member G3 earns 4+17.6=21.6 points.

Information about the contributions to the project

After completion of each contribution decision you receive the following information:

- Average contributions of members of the green group to the green group's project
- Average contributions of members of the blue group to the blue group's project
- Your earnings in the round

Payment

- Upon completion of the 20 decision rounds, you will be informed about your final earnings, the sum of the 20 per-round earnings, on your computer screen.
- All earned points will be converted and paid in cash according to the following exchange rate: **10 points = 0.25 Euro.**

- Afterwards, every participant will be called separately and earnings will be paid in private and anonymously. You will not find out about the earnings of other participants; likewise, other participants will not find out about your earnings.

2.B Post-experimental questionnaire

[*Open questions*]

1. Please explain the motives of your investment decisions in the experiment.
2. Did the contributions of your group differ from the contributions of the other group? If the answer is yes, how do you explain these differences?

[*Behavioral motivations*]

Please indicate to what extent the following motives played a role in your investment decisions. Please read the statements thoroughly and indicate to what extent you agree or disagree. You choose between 1 “I strongly disagree" and 7 “I agree absolutely".

[*Seven-point scales were used for each statement:*]

I strongly disagree [ ] [ ] [ ] [ ] [ ] [ ] [ ] I strongly agree

1. I wanted to earn as much as possible, independent of what others in my group were earning.
2. I wanted to earn more than my group members.
3. I wanted all of my group members to earn as much as possible, regardless of what members of the other group earn.
4. I wanted my group to earn more than the other group.
5. I took the investment decisions of the other group as an orientation for my own investments.
6. I interpreted the situation as a competition between groups.

[*Group identity elicitation*]

1. I feel connected to the other members of my group.
2. I identify as a member of my group.
3. I am happy to be part of my group.

[*Interaction situation*]

1. I was angry when my group contributed less to the group project than members of the other group did to their group project.

2. I was happy when my group contributed less to the group project than members of the other group did to their group project.

3. I was indifferent about whether my group contributed less to the group project than members of the other group did to their group project.

4. The other group has contributed more to its group project than my own group has to its group project.

5. The other group faced better conditions for contributing to its group project than my group.

6. The contributions of the other group are not comparable to the contributions in my group.

7. In my group, the general willingness to contribute was high.

8. In the other group, the general willingness to contribute was high.

[*Demographic questions*]

1. Please indicate your age.

2. Please indicate your gender.

3. If you are a student, please indicate your subject.

2.C Descriptive statistics

Table 2.4 shows an outline of descriptive statistics. Note that Result 2 (LOW), the higher sensitivity to within-group comparisons in LOW-INFO is also evident from this table: The contribution change in response to unfavorable within-group comparisons is particularly large in treatment LOW-INFO, even when we do not account for behavioral patterns known from the literature. Contribution levels are similar across treatments with MPCR $\mu = 0.4$ and show the usual decline in contributions known from the literature.

	No difference		**Structural difference**			
	SAME-NOINFO	SAME-INFO	LOW-NOINFO	HIGH-NOINFO	LOW-INFO	HIGH-INFO
First round	10.5	9.38	9.96	12.96	10.52	11.21
	(7.17)	(6.39)	(6.65)	(6.72)	(7.64)	(6.77)
Round 1-10	7.57	5.89	7.06	11.75	7.33	10.45
	(4.79)	(4.09)	(4.85)	(5.23)	(5.15)	(5.51)
Round 11-20	4.56	2.54	3.36	9.59	3.69	8.72
	(2.72)	(2.23)	(2.88)	(4.29)	(2.99)	(4.77)
Overall	6.07	4.22	5.21	10.67	5.51	9.57
	(3.76)	(3.16)	(3.86)	(4.76)	(4.07)	(5.14)
Contribution	-0.46	-0.47	-0.44	-0.38	-0.50	-0.23
Change	(4.44)	(4.21)	(4.79)	(3.69)	(5.42)	(4.78)
Change: Unfavorable	-2.79	-3.34	-2.83	-2.05	-3.74	-2.02
Comparisons	(1.28)	(1.51)	(1.29)	(1.32)	(1.43)	(1.06)
Change: Favorable	1.31	1.27	1.20	1.93	1.52	1.83
Comparisons	(1.00)	(0.77)	(0.61)	(1.73)	(0.93)	(1.12)

Table 2.4: Group-aggregated means and standard deviations in parentheses by treatment.

2.D Statistical model equations

We use the function *lmer* in R to estimate our models. The matrix form of our models reads,

$$\begin{aligned} \mathrm{y} &= \mathrm{X}\beta + \mathrm{Z}\mathrm{b} + e, \\ \mathrm{b} &\sim \mathcal{N}(0, \Sigma), \\ \Sigma &= \sigma^2 \Lambda_\theta \Lambda_\theta^\top, \end{aligned} \tag{2.1}$$

where y is the n-dimensional dependent variable, n is the number of observations, β is a m-dimensional fixed effects vector, X is an n x m model matrix, which differs on the specified model, and e is the n-dimensional error vector with $e \sim \mathcal{N}(0, \sigma^2 \mathrm{I}_n)$. Z is the n x q model matrix for the q-dimensional random effects variable, Σ_θ is a symmetric variance-covariance matrix generated by the relative covariance factor, Λ_θ, which is a q x q block diagonal matrix depending on the variance component parameter θ. As random effects, we consider nested random intercepts, so the random effects variable is the vector $b = (b_{\mathrm{p}}, b_{\mathrm{pg}}, b_{\mathrm{pgi}})$, where p indicates the pair of groups, g the group, and i the individual. As regards clustering of standard errors, i.e., the variance-covariance matrix of the random intercepts Λ_θ, we do not specify a particular correlation structure, but let the function *lmer* estimate correlations using its default procedure (similar to exchangeable). If we directly specify an explicit variance-covariance structure (e.g., exchangeable, AR(1)) or assume random slopes for the variable *Round*, our significance patterns remain unchanged.

Our dependent variabe y is the *Contribution Change* and defined as $c_{p,g,i,t} - c_{p,g,i,t-1}$, where $c_{p,g,i,t}$ is the contribution in round t of individual i in group g in the pair of groups p. The variable $x_{\text{Unfavorable Comparison}_{pgit}}$, for example, equals 1 if $c_{p,g,i,t-1} - \bar{c}_{p,g,-i,t-1} > 0$ and equals 0 otherwise, where $\bar{c}_{p,g,-i,t-1}$ is the mean contribution of the three other members of i's group in round $t-1$.

For the sake of brevity, we do not write out equations for every estimated model, but focus on L1-L4. Model L1 is

$$\begin{aligned} y_{pgit} =& \beta_0 + x_{\text{Round}_p}\beta_{\text{Round}} + x_{\text{Unfavorable Comparison}_{pgit}}\beta_{\text{Unfavorable Comparison}} + \\ & x_{\text{LOW-NOINFO}_p}\beta_{\text{LOW-NOINFO}} + x_{\text{SAME-INFO}_p}\beta_{\text{SAME-INFO}} + x_{\text{SAME-NOINFO}_p}\beta_{\text{SAME-NOINFO}} + \\ & x_{\text{Group behind Comparison Group}_{pgit}}\beta_{\text{Group behind Comparison Group}} + \\ & z_p b_p + z_{pg} b_{pg} + z_{pgi} b_{pgi} + e_{pgit}. \end{aligned} \tag{2.2}$$

The model L2 is

$$\begin{aligned} y_{pgit} =& \beta_0 + x_{\text{Round}_p}\beta_{\text{Round}} + \\ & x_{\text{LOW-NOINFO}_p}\beta_{\text{LOW-NOINFO}} + \\ & x_{\text{SAME-INFO}_p}\beta_{\text{SAME-INFO}} + \\ & x_{\text{SAME-NOINFO}_p}\beta_{\text{SAME-NOINFO}} + \\ & x_{\text{Unfavorable Comparison}_{pgit}}\beta_{\text{Unfavorable Comparison}} + \\ & x_{\text{Group behind Comparison Group}_{pgit}}\beta_{\text{Group behind Comparison Group}} + \\ & x_{\text{Unfavorable Comparison}_{pgit}} * x_{\text{LOW-NOINFO}_{pg}}\beta_{\text{Unfavorable Comparison*LOW-NOINFO}} + \\ & x_{\text{Unfavorable Comparison}_{pgit}} * x_{\text{SAME-INFO}_p}\beta_{\text{Unfavorable Comparison*SAME-INFO}} + \\ & x_{\text{Unfavorable Comparison}_{pgit}} * x_{\text{SAME-NOINFO}_{pg}}\beta_{\text{Unfavorable Comparison*SAME-NOINFO}} + \\ & z_p b_p + z_{pg} b_{pg} + z_{pgi} b_{pgi} + e_{pgit}. \end{aligned} \tag{2.3}$$

The model L3 is

$$
\begin{aligned}
y_{pgit} = & \beta_0 + x_{\text{Round}_p}\beta_{\text{Round}} + \\
& x_{\text{LOW-NOINFO}_p}\beta_{\text{LOW-NOINFO}} + \\
& x_{\text{SAME-INFO}_p}\beta_{\text{SAME-INFO}} + \\
& x_{\text{SAME-NOINFO}_p}\beta_{\text{SAME-NOINFO}} + \\
& x_{\text{Unfavorable Comparison}_{pgit}}\beta_{\text{Unfavorable Comparison}} + \\
& x_{\text{Favorable Comparison}_{pgit}}\beta_{\text{Favorable Comparison}} + \\
& x_{\text{Group behind Comparison Group}_{pgit}}\beta_{\text{Group behind Comparison Group}} + \\
& x_{\text{Unfavorable Comparison}_{pgit}}x_{\text{LOW-NOINFO}_{pg}}\beta_{\text{Unfavorable Comparison*LOW-NOINFO}} + \\
& x_{\text{Unfavorable Comparison}_{pgit}}x_{\text{SAME-INFO}_p}\beta_{\text{Unfavorable Comparison*SAME-INFO}} + \\
& x_{\text{Unfavorable Comparison}_{pgit}}x_{\text{SAME-NOINFO}_{pg}}\beta_{\text{Unfavorable Comparison*SAME-NOINFO}} + \\
& z_p b_p + z_{pg} b_{pg} + z_{pgi} b_{pgi} + e_{pgit}.
\end{aligned}
\tag{2.4}
$$

The model L4 also includes data of the questionnaire variable $x_{NotComparable}$ which we center at the grand mean, so it equals $x_{\text{raw-Not Comparable}_{pgi}} - \bar{x}_{\text{Not Comparable}}$, where $raw - NotComparable$ is the raw response data in {1,2,...,7}. In model L5, we also center all questionnaire variables at their grand mean.
The model L4 is

$$
\begin{aligned}
y_{pgit} = & \beta_0 + x_{\text{Round}_p}\beta_{\text{Round}} + \\
& x_{\text{LOW-NOINFO}_{pg}}\beta_{\text{LOW-NOINFO}} + \\
& x_{\text{SAME-INFO}_p}\beta_{\text{SAME-INFO}} + \\
& x_{\text{SAME-NOINFO}_p}\beta_{\text{SAME-NOINFO}} + \\
& x_{\text{Unfavorable Comparison}_{pgit}}\beta_{\text{Unfavorable Comparison}} + \\
& x_{\text{Favorable Comparison}_{pgit}}\beta_{\text{Favorable Comparison}} + \\
& x_{\text{Group behind Comparison Group}_{pgit}}\beta_{\text{Group behind Comparison Group}} + \\
& x_{\text{Not Comparable}_{pgi}}\beta_{\text{Not Comparable}} + \\
& x_{\text{Unfav. Comparison}_{pgit}} * x_{\text{LOW-NOINFO}_{pg}}\beta_{\text{Unfav. Comparison*LOW-NOINFO}} + \\
& x_{\text{Unfav. Comparison}_{pgit}} * x_{\text{SAME-INFO}_p}\beta_{\text{Unfav. Comparison*SAME-INFO}} + \\
& x_{\text{Unfav. Comparison}_{pgit}} x_{\text{LOW-NOINFO}_{pg}} x_{\text{Not Comparable}_{pgi}} \\
& \beta_{\text{Unfav. Comparison*LOW-NOINFO*Not Comparable}} + \\
& x_{\text{Unfav. Comparison}_{pgit}} x_{\text{SAME-INFO}_p} x_{\text{Not Comparable}_{pgi}} \\
& \beta_{\text{Unfav. Comparison*SAME-INFO*Not Comparable}} + \\
& x_{\text{Unfav. Comparison}_{pgit}} x_{\text{SAME-NOINFO}_{pg}} x_{\text{Not Comparable}_{pgi}} \\
& \beta_{\text{Unfav. Comparison*SAME-NOINFO*Not Comparable}} + \\
& z_p b_p + z_{pg} b_{pg} + z_{pgi} b_{pgi} + e_{pgit}.
\end{aligned}
\tag{2.5}
$$

2.E Mixed-effect models H0-H4 with MPCR $\mu = 0.8$

We plot the respective models in Table 2.5. For a description of the models, please see the section "Results" in the main text.

	Dependent variable:				
	Contribution change				
	(H0)	(H1)	(H2)	(H3)	(H4)
Round		−0.056***	−0.056***	−0.051**	−0.050**
		(0.020)	(0.020)	(0.020)	(0.020)
Unfavorable Comparison		−3.358****	−3.440****	−2.359****	−2.477****
		(0.220)	(0.309)	(0.449)	(0.477)
HIGH-NOINFO	−0.111	−0.210	−0.282	−0.139	−0.102
	(0.232)	(0.218)	(0.291)	(0.316)	(0.326)
Not Comparable					0.061
					(0.111)
Group behind Comparison Group	0.457*	0.376	0.378	0.314	0.336
	(0.276)	(0.260)	(0.260)	(0.273)	(0.274)
Favorable Comparison				1.667****	1.642****
				(0.374)	(0.378)
Unfavorable Comparison*HIGH-NOINFO			0.165	0.040	0.157
			(0.438)	(0.456)	(0.486)
Unfavorable Comparison*Not Comparable					0.049
					(0.166)
Not Comparable*HIGH-NOINFO					−0.095
					(0.148)
Unfavorable Comparison*Not Comparable*HIGH-NOINFO					0.117
					(0.228)
Constant	−0.359**	1.800****	1.836****	0.556	0.526
	(0.180)	(0.305)	(0.320)	(0.460)	(0.463)
Observations	1,824	1,824	1,824	1,824	1,824
Log Likelihood	−5,498.383	−5,387.234	−5,387.163	−5,377.743	−5,376.648
Akaike Inf. Crit.	11,010.760	10,792.470	10,794.330	10,777.490	10,783.300
Bayesian Inf. Crit.	11,049.330	10,842.050	10,849.410	10,838.080	10,865.930
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 2.5: Mixed-effects models showing Result 1-3 (HIGH) from the main text.

2.F Additional analyses

In the main text, we present mixed-effect models with random intercepts that account for dependencies between pairs of groups, groups, and individuals. In Table 2.6, we extend Model L3 from the main text (replotted here from Table 2.3 for convenience) by either interacting treatment dummies and *Group behind Comparison Group* (Model L3a) or interacting treatment dummies and *Favorable Comparison* (Model L3b). Coefficients for the interactions between *Unfavorable Comparison* and treatment dummies remain qualitatively and quantitatively robust and significantly different from the reference category LOW-INFO, thus corroborating Result 2 in the main text. Note that Model L3b is plagued by multicollinearity, which inflates variance: A maximum variance inflation factor (*max VIF*) larger than 10 is strongly indicative of multicollinearity (here, *max VIF*(Model L3b) = 12.91 $>$ 10, *max VIF*(Model L3) and *max VIF*(Model L3a) $<$ 10). We extend Models L1-L4 and H1-H4 from the main text by the variable *Comparison Group Minus Group* in Table 2.7 and Table 2.8. These extensions show that how far a group is behind the comparison group is an important predictor of contributions, but leaves the significance patterns of other variables virtually unchanged.

	Dependent variable:		
	Contribution Change		
	(L3)	(L3a)	(L3b)
Round	−0.023**	−0.025**	−0.023**
	(0.011)	(0.011)	(0.011)
LOW-NOINFO	−0.377	−0.484	−0.375
	(0.263)	(0.471)	(0.565)
SAME-INFO	−0.242	0.204	−0.395
	(0.227)	(0.352)	(0.434)
SAME-NOINFO	−0.238	0.191	−0.378
	(0.228)	(0.353)	(0.439)
Unfavorable Comparison	−4.224***	−4.172***	−4.332***
	(0.331)	(0.333)	(0.440)
Group behind Comparison Group	0.265**	0.714**	0.265**
	(0.127)	(0.326)	(0.127)
Favorable Comparison	0.906***	0.905***	0.761*
	(0.176)	(0.176)	(0.426)
Unfavorable Comparison*LOW-NOINFO	1.195***	1.188***	1.194*
	(0.423)	(0.426)	(0.655)
Unfavorable Comparison*SAME-INFO	0.855**	0.767**	1.008*
	(0.369)	(0.372)	(0.522)
Unfavorable Comparison*SAME-NOINFO	1.052***	0.976***	1.191**
	(0.367)	(0.370)	(0.525)
Group behind Comparison Group*LOW-NOINFO		0.090	
		(0.488)	
Group behind Comparison Group*SAME-INFO		−0.647*	
		(0.387)	
Group behind Comparison Group*SAME-NOINFO		−0.612	
		(0.386)	
Favorable Comparison*LOW-NOINFO			0.010
			(0.639)
Favorable Comparison*SAME-INFO			0.210
			(0.506)
Favorable Comparison*SAME-NOINFO			0.190
			(0.511)
Constant	0.727**	0.403	0.837**
	(0.302)	(0.373)	(0.418)
Observations[individuals|groups|pairs of groups]	5,472[288|72|48]	5,472[288|72|48]	5,472[288|72|48]
Log Likelihood	−15,852	−15,849	−15,852
Akaike Inf. Crit.	31,734	31,734	31,739
Bayesian Inf. Crit.	31,833	31,853	31,858
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 2.6: Additional mixed-effects models showing the robustness of Result 2 (LOW) from the main text.

	Dependent variable:			
	Contribution Change			
	(L1')	(L2')	(L3')	(L4')
Round	−0.064****	−0.064****	−0.038***	−0.039***
	(0.013)	(0.013)	(0.012)	(0.012)
Unfavorable Comparison	−4.118****	−4.961****	−4.102****	−3.773****
	(0.127)	(0.305)	(0.333)	(0.387)
LOW-NOINFO	0.140	−0.293	−0.337	−0.444
	(0.265)	(0.315)	(0.262)	(0.319)
SAME-INFO	−0.182	−0.500*	−0.368	−0.449
	(0.242)	(0.280)	(0.231)	(0.287)
SAME-NOINFO	0.048	−0.359	−0.279	−0.336
	(0.234)	(0.275)	(0.228)	(0.284)
Not Comparable				−0.034
				(0.108)
Favorable Comparison			0.950****	0.938****
			(0.176)	(0.177)
Group behind Comparison Group	0.364***	0.372***	0.297**	0.311**
	(0.139)	(0.140)	(0.127)	(0.127)
Comparison Group Minus Group	0.038****	0.038****	0.023***	0.024***
	(0.010)	(0.010)	(0.008)	(0.008)
Unfavorable Comparison*LOW-NOINFO		1.133***	1.130***	0.924*
		(0.426)	(0.423)	(0.484)
Unfavorable Comparison*SAME-INFO		0.854**	0.802**	0.559
		(0.373)	(0.368)	(0.419)
Unfavorable Comparison*SAME-NOINFO		1.085***	1.002***	0.661
		(0.371)	(0.367)	(0.418)
Unfavorable Comparison*Not Comparable				−0.420***
				(0.157)
LOW-NOINFO*Not Comparable				−0.046
				(0.135)
SAME-INFO*Not Comparable				−0.049
				(0.125)
SAME-NOINFO*Not Comparable				0.022
				(0.127)
Unfavorable Comparison*LOW-NOINFO*Not Comparable				0.571***
				(0.210)
Unfavorable Comparison*SAME-INFO*Not Comparable				0.710****
				(0.189)
Unfavorable Comparison*SAME-NOINFO*Not Comparable				0.409**
				(0.190)
Constant	2.071****	2.378****	1.159****	1.239***
	(0.333)	(0.349)	(0.339)	(0.377)
Observations	5,472	5,472	5,472	5,472
Log Likelihood	−15,864.860	−15,859.910	−15,848.010	−15,835.500
Akaike Inf. Crit.	31,753.710	31,749.810	31,728.020	31,718.990
Bayesian Inf. Crit.	31,833.000	31,848.930	31,833.740	31,877.570
Note:				*p<0.1; **p<0.05; ***p<0.01; ****p<0.001

Table 2.7: Mixed-effects models L1'-L4' from Result 1 (LOW) from the main text.

	Dependent variable:			
	Contribution Change			
	(H1')	(H2')	(H3')	(H4')
Round	−0.112****	−0.112****	−0.078****	−0.078****
	(0.023)	(0.023)	(0.022)	(0.022)
Unfavorable Comparison	−4.020****	−4.102****	−2.882****	−3.007****
	(0.238)	(0.336)	(0.497)	(0.528)
HIGH-NOINFO	−0.121	−0.186	−0.132	−0.092
	(0.375)	(0.429)	(0.350)	(0.360)
Not Comparable				0.065
				(0.123)
Favorable Comparison			1.475****	1.440****
			(0.431)	(0.434)
Group behind Comparison Group	0.715**	0.716**	0.397	0.410
	(0.311)	(0.310)	(0.286)	(0.287)
Comparison Group Minus Group	0.094****	0.093****	0.043**	0.043**
	(0.019)	(0.019)	(0.017)	(0.017)
Unfavorable Comparison*HIGH-NOINFO		0.148	0.075	0.205
		(0.476)	(0.476)	(0.506)
Unfavorable Comparison*Not Comparable				0.063
				(0.176)
HIGH-NOINFO*Not Comparable				−0.121
				(0.165)
Unfavorable Comparison*HIGH-NOINFO*Not Comparable				0.123
				(0.241)
Constant	4.075****	4.104****	1.826**	1.802**
	(0.565)	(0.571)	(0.711)	(0.720)
Observations	1,824	1,824	1,824	1,824
Log Likelihood	−5,380.368	−5,380.321	−5,375.202	−5,374.062
Akaike Inf. Crit.	10,780.740	10,782.640	10,774.400	10,780.120
Bayesian Inf. Crit.	10,835.820	10,843.240	10,840.510	10,868.260
Note:				*p<0.1; **p<0.05; ***p<0.01; ****p<0.001

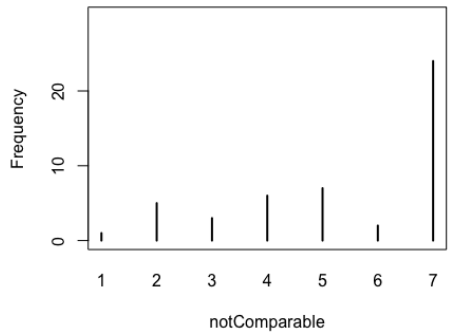
Table 2.8: Mixed-effects models H1'-H4' from Result 1 (HIGH) from the main text.

2.G Questionnaire data

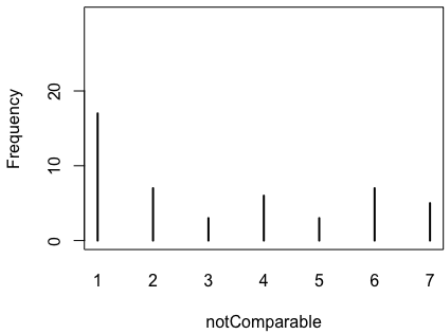
	No difference		Structural difference			
	SAME-NOINFO	SAME-INFO	LOW-NOINFO	HIGH-NOINFO	LOW-INFO	HIGH-INFO
Group Comparison [BM5]	2.8	2.7	2.9	3.1	2.8	2.8
Identity [GI1-3]	3.2	3.1	2.6	2.4	4.2	4.3
Anger [IS1]	4.2	4.2	4.6	3.8	3.4	3.5
Joy [IS2]	4.6	4.6	3.8	4.2	5.4	4.4
Not Comparable [IS6]	3.8	3.8	3.3	4.0	5.4	5.0
OtherBetterCond [IS7]	2.4	2.8	2.7	2.0	6.0	2.0
OwnCoop [IS8]	3.2	2.8	2.1	4.7	2.3	4.5

Table 2.9: Means for the most interesting questionnaire variables by treatment (Likert-Scale {1,2,..,7}). BM5 abbreviates "Behavioral Motivations, Question 5", analogously GI to the "Group identity elicitation" and IS refers to questions about the "Interaction situation".

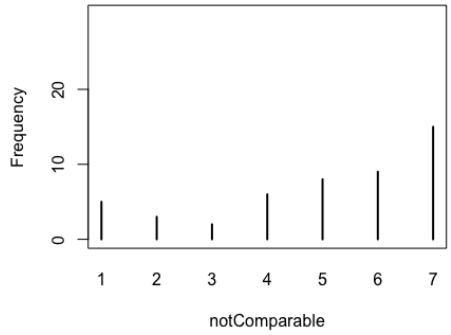
Table 2.9 shows that perceptions of the interaction situation differ considerably between treatments. Our elicitation of group identity differs largely depending on whether the own group has a high MPCR or not. In addition, individuals in LOW-INFO seem to be less angry when their group lags behind a comparison (variable *Anger*), but also happier when their group lies ahead (variable *Joy*). Presumably this is because individuals have information about the comparison group' s 'better' conditions (as can also be seen from the variable *OtherBetterCond*). The perception of whether groups are not comparable differs between LOW-INFO and LOW-NOINFO. Perhaps more intriguingly, looking the histograms in Figure S2.4, we see that *NotComparable* even differs between LOW-INFO and HIGH-INFO, suggesting that this perception might be partly self-serving.



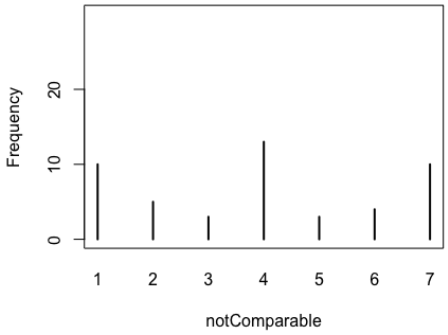
(a) LOW-INFO.



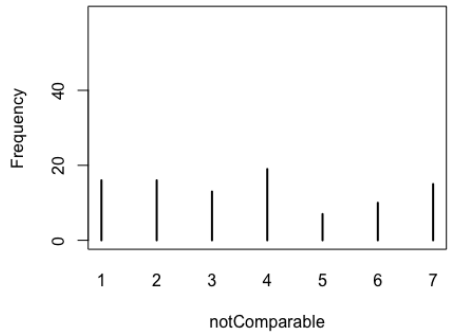
(b) LOW-NOINFO.



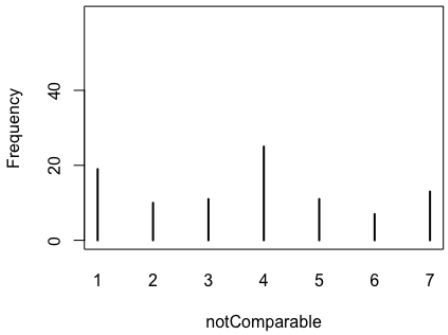
(c) HIGH-INFO.



(d) HIGH-NOINFO.



(e) SAME-INFO.



(f) SAME-NOINFO.

Figure 2.4: The absolute frequencies of different levels of agreement {1,2,..,7} to the statement "The contributions of the other group are not comparable to the contributions in my group".

Chapter 3

The driving forces behind information acquisition in social decisions

3.1 Introduction

In times of mass media, social media, and digitalization, we are surrounded by multiple sources of information. People may not only decide which information to follow, but also which information to acquire in the first place. Consider, as an example, the phenomenon of climate change denial. Although scientifically-acknowledged evidence on climate change is available, many people are reluctant to become involved with this information, regardless of their ability to understand scientific facts (Kahan et al., 2012) or hints on social media (Bakshy et al., 2015). Previous accounts explain such information avoidance by people's desire to maintain positive beliefs about themselves (Bem, 1972; Mazar et al., 2008; Grossman and van der Weele, 2017), e.g., that they would have acted differently had they known better. Experiments have shown that people avoid information which might confront them with the potentially adverse effects of their own actions (Kuang et al., 2007; Larson and Capra, 2009; Conrads and Irlenbusch, 2013; Grossman, 2014). When avoiding certain information, people also prevent themselves from being able to inform others. For example, sociological in-depth interviews in a Norwegian fishery village suggest that although most people are aware of climate change, they hesitate to gather further information about the issue and mention it in front of others, partly to manage how others think of them (Norgaard, 2006). People might thus avoid information more strongly if it affects not only what they think of themselves but also what others think of them (Bénabou and Tirole, 2006; Battigalli and Dufwenberg, 2007; Andreoni and Bernheim, 2009).

In this paper, we focus on information acquisition in the context of prosocial behavior. Empirical evidence suggests that business leaders and politicians sometimes pretend ignorance as an excuse for their anti-social behavior (McGraw, 1991; Luban, 1999; Kadish et al., 2016). Similarly, evidence from the lab shows that ignorant individuals receive less punishment for a given anti-social behavior (Bartling et al., 2014).[1] However, the drivers of information acquisition in the context of prosocial behavior are not yet fully understood. We therefore investigate systematically how people acquire information (i) if it affects their prosocial appearance in front of others and (ii) if there are monetary incentives to appear prosocial. To this end, we use an extended dictator game. Dictators decide on a transfer from their endowment to an anonymous recipient. Prior to deciding, dictators can acquire social comparison information, i.e., information about the transfers of previous dictators.[2] In particular, they can choose between transfer information from previous dictators with an equally high endowment or from dictators with a lower (higher) than their own endowment. The dictator may expect that the previous dictator with a lower (higher) endowment made a lower (higher) transfer than the one with the same endowment. A dictator who takes the previous dictator as a reference for the own transfer (in absolute terms) may thus strategically choose the comparison information. As a first treatment variation, we vary whether the acquired transfer information is private or whether the transfer information (but not the endowment of the previous dictator) is made public. This allows to disentangle information acquisition when it is merely motivated by affecting the own fairness perceptions and when it is also motivated by affecting the fairness perceptions of others and thus the own prosocial appearance. As a second treatment variation, we vary whether dictators can be punished by recipients. This allows investigating how information acquisition changes when we increase the (monetary) incentives for appearing prosocial.

We find that the share of individuals who avoid information from previous dictators with an equally high or higher endowments and rather acquire information from previous dictators with a lower endowment increases by more than 50% (from 23% to 38% on average) when the acquired transfer information is made public and no longer private. Interestingly, the information

[1] More generally, people often excuse low prosocial behavior due to information about external circumstances (Linardi and McConnell, 2011), risks in giving (Haisley and Weber, 2010; Exley, 2016), their responsibility (Bartling and Fischbacher, 2012; Falk and Szech, 2014; Bartling et al., 2015) or the 'worthiness' or race of donation recipients (Fong and Luttmer, 2011; Fong and Oberholzer-Gee, 2011; Spiekermann and Weiss, 2016). If possible, many people also avoid demands for prosocial behavior (Dana et al., 2006; Dellavigna et al., 2012; Damgaard and Gravert, 2013; Knutsson et al., 2013; Adena and Huck, 2016; Andreoni et al., 2017).

[2] A broad literature now demonstrates that people are more willing to transfer money to anonymous others if others do so, too (Cason and Mui, 1998; Frey and Meier, 2004; Bicchieri and Xiao, 2009; Krupka and Weber, 2009; Shang and Croson, 2009).

avoidance we observe is independent of whether or not recipients can punish dictators. Hence, we provide evidence that the desire to appear prosocial to others motivates information avoidance, but less so increased monetary incentives to appear prosocial.

As an additional contribution, our paper sheds light on 'sorting', i.e., about what the act of information acquisition reveals about an individual's prosocial motivation and behavior. If there is 'sorting' by information acquisition, those who avoid information suggesting high transfers should also be less prosocially motivated *per se*, i.e., they would also make lower transfers than non-avoiders had they received *the same* information. If, by contrast, there is no sorting, those who avoid information suggesting high transfers make lower transfers merely due to the different information they receive. A better understanding of sorting is important to find out whether, e.g., climate change denial predicts individual differences in environmental concerns and behavior, or whether it simply reflects a person's lack of information. Former laboratory studies have investigated sorting (Fong and Oberholzer-Gee, 2011; Conrads and Irlenbusch, 2013; Grossman and van der Weele, 2017), but did not find clear-cut results. Importantly, the design of previous studies did not allow observing whether or not avoiders would have exhibited different behavior, were they exposed to different information. Our information acquisition procedure (see Experimental Design for details) allows for such observations. We present first clean evidence on 'sorting', namely that information avoidance predicts lower prosocial behavior and thus heterogeneity in prosocial motivation, independent of the received information.

3.2 Experiment

3.2.1 Participants and Procedure

We recruited our participants via the MTurk US online labor market.[3] Subjects participated in an extended online dictator game. We programmed the experiment using the software SosciSurvey. We collected all data during weekdays in May and June 2016 and randomly assigned treatments to different days. The software randomly assigned participants to either the role of the 'dictator' or the 'recipient' (Player A or B in the instructions), so dictators and recipients arrived at our game using the same advert on MTurk. This procedure ensures credibility of the study as

[3]MTurk workers pay as much or more attention to instructions compared to regular student subject pools unless experiments last very long (Goodman et al., 2013; Hauser and Schwarz, 2016). Many classic experimental economics findings have been replicated on MTurk (Horton et al., 2011; Suri and Watts, 2011), and online dictator games have proven particularly insightful (Fong and Luttmer, 2011; Amir et al., 2012; Ockenfels and Werner, 2012; Raihani et al., 2013; Raihani and McAuliffe, 2014; Jordan et al., 2015).

opposed to an ex-post matching procedure, where MTurk workers might find out that, for a given advert, everybody receives the role of the dictator. Instructions ensured the anonymity of MTurk workers. Participants earned a \$1 participation fee solely for taking part in the experiment. After completion of the game, participants filled in a questionnaire, asking for age, gender, and experience with similar games. A message accompanying the payment of additional payoffs informed participants about transfer and punishment decisions (where applicable). The average payoff earned during the game in addition to the participation fee was \$1.85 for dictators and \$1.01 for recipients. Stakes in our online experiment were similar to laboratory experiments. Dictators took between 5 to 10 minutes to complete the experiment, resulting in a minimum hourly wage of \$17.11 (including participation fee). Altogether, our data base consists of the decisions of 411 independent pairs of dictators and recipients.[4]

3.2.2 Experimental Design

Participants find out whether they have been assigned the role of dictator or the role of recipient before the game starts. Dictators receive an endowment of \$2. As in a standard dictator game, they can transfer a share of this endowment to a passive recipient, that is, choose a transfer in [\$0.00, \$0.01, ... , \$2.00]. Different from a standard dictator game, only the dictator knows her precise endowment, but not the recipient. Recipients only know that the dictator has an endowment of either \$1, \$2, or \$4, and that dictators know about this information asymmetry. We introduce this information asymmetry, such that recipients cannot infer the relative share implied by their dictator's transfer. Due to this ambiguity, information about what is appropriate is likely to be more important.[5] Dictators proceed through two decision stages: the information acquisition stage and the transfer stage.

In the *information acquisition stage*, dictators choose among multiple sources of social comparison information. We designed our information choice procedure to allow inferences on the motivation behind choices. In the information acquisition stage, the dictators have to choose from which previous dictator, hence-forth *comparison dictator*, they will receive information about his or her transfer. In *Choice 1*, dictators choose between transfer

[4]Of the 1079 recruited participants 153 dropped out and 100 were excluded by the program because they did not answer the control questions correctly within three trials. The share of participants who dropped out or were excluded did not vary systematically across treatments (Chi-Square-Test, $X^2 = 1.67$, $p = 0.643$). These numbers are similar to other MTurk studies (e.g. Jordan et al., 2015). In addition, we had to exclude 4 observations due to an error in the matching procedure (three recipients were accidentally matched to the same dictator).

[5]See Ockenfels and Werner (2012) for a dictator game with asymmetric information about the stake size and Mitzkewitz and Nagel (1993) for an ultimatum game. Dictators often allocate an equal split of the lowest possible endowment next to an equal split of their own endowment.

information of a comparison dictator with a *low* endowment ($ 1) and transfer information of a comparison dictator with a *medium* endowment ($ 2, equal to their own endowment).[6] In *Choice 2*, dictators choose between transfer information of a comparison dictator with a *medium* endowment ($2) and transfer information of a comparison dictator with a *high* endowment ($4). Importantly, in each of the two choices, dictators can choose to receive transfer information of a comparison dictator with the same endowment as their own ($2) to find out what others did when they were in the same situation as they are (*diagnostic information search*). In Choice 1, the alternative to the diagnostic information is a dictator with a low endowment. Choosing this information may be motivated by appearing prosocial, as it is more likely to result in a lower transfer information for comparison. In Choice 2, the alternative is a dictator with a *high* endowment. Subjects motivated to appear prosocial may avoid this alternative, as it is likely to result in a higher transfer information for comparison.[7] Thus, dictators' decisions in the two choice tasks results in one of the four possible (combinations of) sources of information: *low*($1)-*medium*($2), *medium*($2)-*medium*($2), *medium*($2)-*high*($4), and *low*($1)-*high*($4). Choosing *low-medium* results in receiving transfer information from comparison dictators with the lowest possible endowments in both choices, and thus most likely contains information suggesting low prosocial behavior. By contrast, choosing *medium-medium* is most likely associated with diagnostic motives to update the own fairness perceptions. After dictators have decided on Choice 1 and Choice 2, one of their two choices is randomly determined to be relevant. Dictators then receive the transfer information of the chosen comparison dictator. In the subsequent *transfer decision stage*, dictators decide on their transfer to their recipient.[8]

Implementing a procedure with two choice tasks and randomly determining which task becomes relevant has two major advantages. First, it yields a greater statistical power to distinguish motivated choices from purely random choice of a participant not at all interested in social comparisons (standard economic preferences). Second, the procedure generates observations of dictators who differ in their information choice, but nonetheless receive the same transfer information. For example, dictators choosing *low-medium, medium-medium*, or

[6] We randomized the order of the alternatives for each of the two choices.

[7] Selecting *high* may also have the diagnostic content of upward comparisons. That is, people may perceive others with a higher endowment as upper benchmarks for their own behavior. This interpretation is in line with a broad psychological literature which associates upward comparisons with diagnostic, self-assessment motives (Taylor et al., 1996; Corcoran et al., 2011).

[8] The transfer information reveals either an equal split of the comparison dictator's endowment or a transfer of zero, both with equal probability. We collected the transfer decisions of comparison dictators in a previous experiment, where recipients faced the same information asymmetry regarding the dictator's endowment.

medium-high may all receive transfer information of a comparison dictator with *medium* endowment with some probability. This allows to test whether dictators who chose different sources, but receive the same information, differ in how they follow this information. In particular, we can investigate, for instance, whether dictators choosing *low-medium* and dictators choosing *medium-medium* transfer different amounts, when both receive the same transfer information of a comparison dictator with *medium* endowment. This procedure allows investigating 'sorting', i.e., studying whether dictators who seek information about low transfers act less prosocially *per se*, regardless of the information they actually receive.

3.2.3 Treatments

Our four treatments systematically vary how the acquired information affects fairness perceptions (and thus prosocial appearance) and payoffs. In all four treatments, the recipient knows that the own dictator has an endowment of either \$1, \$2, or \$4. It is also common to all treatments that the recipient does not know that the dictator makes information acquisition choices. In treatment PRIVATE-BASE, only the dictator - and not the recipient - receives the transfer information of the comparison dictator. The recipient only learns how much money the own dictator transfers. This treatment introduces self-image motives in information acquisition by informing only the dictator about how the own transfer compares to that of another individual.

In the PUBLIC-BASE treatment, both the dictator and the recipient receive the transfer information of the comparison dictator. Hence, the revealed information becomes public and can affect not only the dictator's belief about what is appropriate (as in PRIVATE-BASE), but also the recipient's belief. As recipients are unaware that the dictators made information acquisition choices, they cannot infer motives from the received comparison information. Furthermore, recipients can only evaluate the transfer information in absolute terms because they only know that the endowment of the comparison dictator is either \$1, \$2, or \$4, just like they only know that the endowment of their own dictator is either \$1, \$2, or \$4. This treatment additionally makes the social information relevant for the recipient's fairness perceptions and thus for the dictator's prosocial appearance.

In an additional treatment dimension, we provide recipients with the opportunity to engage in costly punishment of their dictator to investigate how the desire to appear prosocial to avoid negative payoff consequences affects the dictators' information choices. In PUBLIC-PUN recipients receive the transfer information of the comparison dictator, as in PUBLIC-BASE, and they find out how much their own dictator has transferred. Afterwards, they receive an

additional endowment of \$1, which they can keep or invest to punish their own dictator. For each cent invested in punishment, three cents are deducted from the dictator's payoff.[9] Dictators also receive an additional endowment of \$1 to hold inequity considerations fixed across treatments.

To complete our 2x2 experimental design, treatment PRIVATE-PUN is equivalent to PUBLIC-PUN, except that recipients do not find out that dictators receive transfer information of a comparison dictator. This treatment allows identifying differences in dictators' transfers between PRIVATE-PUN and PUBLIC-PUN. Further, different from PUBLIC-PUN, dictators can seek information to assess what is appropriate without affecting the recipient's punishment decisions. We summarize the main features of the four treatments in Table 3.1.

	Recipient learns about the dictator's...			***Recipients can punish dictators***	***Dictator is exposed to***	
	possible endowments \$1, \$2, or \$4	choice btw. sources of information	transfer information for comparison		information affecting appearance to recipient	monetary incentives to appear prosocial
PRIVATE-BASE	yes	no	no	no	no	no
PUBLIC-BASE	yes	no	yes	no	yes	no
PRIVATE-PUN	yes	no	no	yes	no	yes
PUBLIC-PUN	yes	no	yes	yes	yes	yes

Table 3.1: Overview of experimental treatments in our 2x2 design.

3.3 Hypotheses

We deduce our hypotheses from self- or social signaling models where agents signal their prosocial motivation via their actions (Bénabou and Tirole, 2006; Andreoni and Bernheim, 2009; Grossman and van der Weele, 2017). We demonstrate the existence of an equilibrium, where, if agents update their fairness perceptions due to previous dictator's transfers, their information acquisition will be indicative of their prosocial motivation. Among the people who care about fairness, the more prosocially motivated types will acquire information suggesting high prosocial behavior and choose high transfers, whereas the less prosocially motivated types will acquire information suggesting low prosocial behavior and choose low transfers. We show that, if social information is made public (as in our PUBLIC treatments) and it thus more strongly affects fairness perceptions than if it is private information, people can enhance their

[9] We labelled units invested in punishment Deduction Cents in the instructions.

prosocial appearance for a given action by acquiring information suggesting low prosocial behavior. If a punishment option is available to recipients (PUN treatments), by contrast, agents can enhance their prosocial appearance only by acquiring information suggesting high prosocial behavior and engaging in high prosocial behavior.

3.3.1 The Model

Following closely the models of Andreoni and Bernheim (2009) and Bénabou and Tirole (2006), we assume that people care about being prosocial and about being perceived as prosocial. Agents are prosocially motivated to the extent v, i.e., which is their social type drawn from a distribution $g(v)$ and not known to an observer. Further, the agents all equally care about their prosocial appearance $E(v|a)$ to the extent μ, i.e., their expected social type inferred by an observer (e.g., the dictator's recipient or an imaginary spectator) who can only access the information i and their action a. Before agents choose their action, they can choose between sources of social information $i \in \{i_L,i_H\}$, where i_L is information suggesting a low transfer and i_H is information suggesting a high transfer, i.e., the normative action. A prosocial action a comes at cost ca and agents choose between a low transfer, i.e., a_L, and a high transfer, the normative action a_H. Following Andreoni and Bernheim (2009), we assume that fairness perceptions f(a,i) depend not only on people's action, a, but also on the distance of their own action to the fairness norm a_H, say a - a_H. We extend their model by assuming that, in addition, fairness perceptions f(a,i) also depend on the distance of their own action to that of comparison standards, a – i. People weigh the impact of the acquired comparison standard i on fairness with p and the impact of the (injunctive) fairness norm i_H on fairness with $1-p$, where $p < 1/2$.[10] The preferences of the decision-maker are $U(a,i) = v$ f(a,i) + μE($v|$ a,i) – ca.

Based on this model, we can derive our hypotheses. We show in the Appendix 3.A that there is a separating equilibrium with threshold v^* . In this equilibrium, agents below the threshold, i.e., the less prosocially motivated agents, acquire information suggesting low transfers i_L and choose a low transfer a_L, whereas the more social types acquire information suggesting high transfers i_H and behave accordingly by choosing a_H. [11]

[10]The resulting fairness perceptions are similar to models of guilt aversion Battigalli and Dufwenberg (2007), where recipients' expectations (and thus guilt) depends on comparison standards and a fairness norm. Preferences involving payoff comparisons also assume that comparison standards affect utility (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000).

[11]Since individuals choosing sources suggesting low transfers i_L have a lower prosocial motivation v than those choosing sources suggesting high transfers i_H, they should also make lower transfers than those choosing sources suggesting high transfers, if only information suggesting high transfers were available.

Hypothesis 1: *Individuals with low prosocial motivation choose a source of information suggesting low transfers and make lower transfers than individuals choosing sources of information suggesting high transfers.*

In the treatments PUBLIC-BASE and PUBLIC-PUN, social information affects not only the own but also other people's fairness perceptions. In particular, making social information public makes the social information relevant for the recipient's fairness perceptions and will thus increase the impact of social information on the agent's fairness perceptions f(a,i), i.e., increase p. An increase in p leads to an increased equilibrium threshold v^* and thus a higher share of agents acquiring information suggesting low transfers i_L. This is because, in equilibrium, the image of making a low transfer E($v|a = a_L$, $i = i_L$) becomes more favorable as the social information makes low transfers relatively less unfair, i.e., f($a = a_L$, $i = i_L$) > f($a = a_L$, $i = i_H$).

Hypothesis 2: *Dictators more often choose a source of information suggesting low transfers when social information is made public, i.e., in treatments PUBLIC-BASE and PUBLIC-PUN, compared to the private information treatments PRIVATE-BASE and PRIVATE-PUN, respectively.*

The availability of a punishment option by the recipient may increase the incentive to appear prosocial, i.e., increase μ. In equilibrium, an increase in μ leads to a decrease in the threshold v^* and thus a lower share of agents acquiring information suggesting low prosocial behavior. This is because, in equilibrium, the prosocial appearance of making a high transfer is better than the prosocial appearance of making a low transfer E($v|a = a_H$,$i = i_H$) > E($v|a = a_L$, $i = i_L$).

Hypothesis 3: *The availability of a punishment option is expected to decrease the number of dictators choosing information suggesting low prosocial behavior, i.e., in treatments* PRIVATE-PUN and PUBLIC-PUN*, compared to treatments PRIVATE-BASE and PUBLIC-BASE, respectively.*

3.4 Results

Information Choice

As a first step, we examine which sources of social comparison information dictators acquire. If dictators were indifferent and chose sources of information at random, they should choose each of the four different combinations with equal probability of 25% . In contrast, a significant deviation from a share of 25% provides evidence for systematic choice of the corresponding

source. In all treatments, we observe choice behavior different from random choice (Chi-Square tests, all p-values < 0.001, $n = 105$ in PRIVATE-BASE, $n = 103$ in PUBLIC-BASE, $n = 102$ in PUBLIC-PUN, $n = 101$ in PRIVATE-PUN).[12]

Dictators can choose the source *low-medium* to more likely receive information about lower transfers of their comparison dictator, which make own transfers appear more prosocial. In line with Hypothesis 2, Figure 3.1A shows that when the received transfer information of the comparison dictator becomes public information (PUBLIC-BASE and PUBLIC-PUN), close to 40% of dictators choose the *low-medium* source, which is significantly more frequent than suggested by chance (Binomial tests: 38 of 103 in PUBLIC-BASE, $p = 0.008$ and 40 of 102 in PUBLIC-PUN, $p = 0.002$). In contrast, when recipients do not receive the transfer information of the comparison dictator, the number of dictators choosing *low-medium* does not exceed the chance level (Binomial tests: 27 of 105 in PRIVATE-BASE, $p = 0.910$ and 20 of 101 in PRIVATE-PUN, $p = 0.252$). In direct support of Hypothesis 2, the number of *low-medium* choices significantly increases under public information as compared to private information settings (PRIVATE-PUN vs. PUBLIC-PUN, Chi-Square test, $X^2 = 9.19$, $p= 0.002$; PRIVATE-BASE vs. PUBLIC-BASE, Chi-Square test, $X^2 = 3.02$, $p = 0.082$).[13] Result 1 sums up the evidence for Hypothesis 2.

Result 1: *Dictators systematically choose the source of information suggesting low prosocial behavior (low-medium) when the acquired information is made public. In both treatments, PUBLIC-BASE and PUBLIC-PUN, a similarly high percentage (close to 40%) chooses low-medium.*

In partial support of Hypothesis 3, the number of *low-medium* choices does not increase significantly when recipients have a punishment opportunity, both under private and public information (PRIVATE-BASE vs. PRIVATE-PUN, Chi-Square test, $X^2 = 1.02$, $p = 0.312$; PUBLIC-BASE vs. PUBLIC-PUN, Chi-Square test, $X^2 = 0.12$, $p = 0.732$), but increases only relative to the baseline (PRIVATE-BASE vs. PUBLIC-PUN, Chi-Square test, $X^2 = 4.30$, $p = 0.038$).

[12] Demographics are well-balanced across treatments (see Appendix 3.B).

[13] See also Appendix Table 3.3, Model 2 (logit regression with demographic controls), $b_{\text{PUBLIC-PUN vs. PRIVATE-PUN}} = 1.00$, SE $= 0.33$, $p = 0.003$ and $b_{\text{PUBLIC-BASE vs. PRIVATE-BASE}} = 0.64$, SE $= 0.32$, $p = 0.044$, respectively.

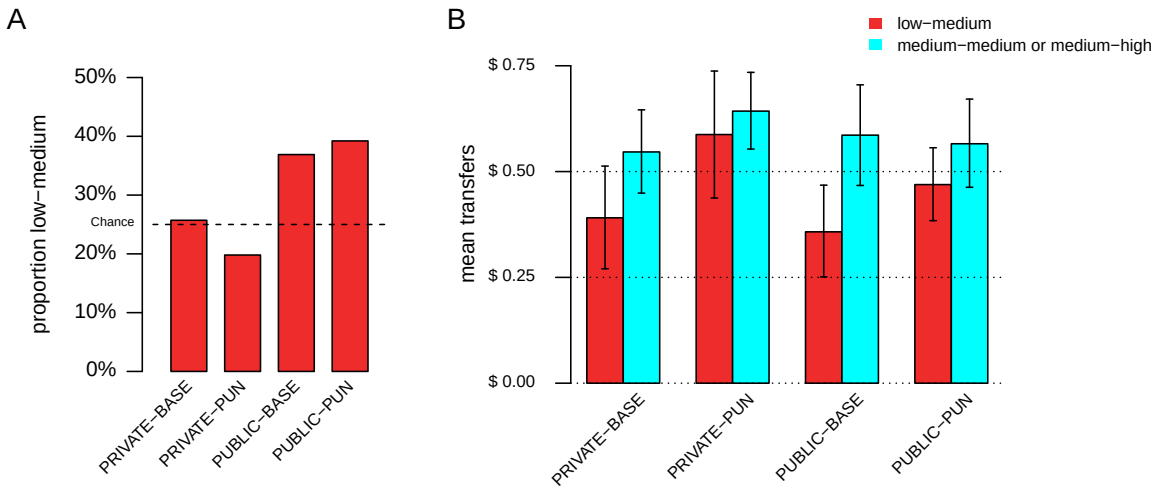


Figure 3.1: **A** Proportion of dictators that chose information from previous dictators with low endowment in Choice 1 and medium endowment in Choice 2 (low-medium), by treatment. **B** Mean transfers with bootstrapped 95% confidence intervals for each treatment, shown separately for dictators choosing information from previous dictators with low endowment in Choice 1 and medium endowment in Choice 2 (*low-medium*) or equally high or higher endowments (*medium-medium* or *medium-high*).

Result 2: *Dictators do not choose the source of information suggesting low prosocial behavior (low-medium) more often when a punishment option is introduced (PRIVATE-PUN vs. PRIVATE-BASE, PUBLIC-PUN vs. PUBLIC-BASE).*

Dictators also choose the source *medium-medium* more often than suggested by chance (25%) in three treatments (Binomial tests: 44 of 105 in PRIVATE-BASE, $p < 0.001$, 41 of 103 in PUBLIC-BASE, $p < 0.001$; 45 of 101 in PRIVATE-PUN, $p < 0.001$). Considerably fewer participants (about one quarter) also choose *medium-high* but this is never different from chance (27 of 105 in PRIVATE-BASE, 20 of 103 in PUBLIC-BASE; 32 of 101 in PRIVATE-PUN, and 27 of 102 in PUBLIC-PUN, Binomial tests, all p-values above 0.135). The sources *medium-medium* and *medium-high* may be associated with diagnostic motives to find out what is appropriate. Only a few participants choose *low-high* (5% in all treatments; 6 of 105 in PRIVATE-BASE, 4 of 103 in PUBLIC-BASE; 4 of 101 in PRIVATE-PUN, and 8 of 102 in PUBLIC-PUN). For this reason, we restrict our attention to the other three information choices.

Transfers

Next, we examine whether dictator transfers vary by information choice. Since transfers do not differ significantly between dictators choosing the sources *medium-medium* and *medium-high* (pooled, \$0.61 vs. \$0.56, Z = -0.76; p = 0.438, for each treatment $p > 0.169$), we pool those observations.[14] In support of Hypothesis 1, dictators who choose information suggesting low prosocial behavior, i.e., *low-medium*, transfer about 26% less than dictators choosing *medium-medium* or *medium-high* (mean \$0.44 vs. \$0.59, Permutation test, Z = 2.89, p = 0.004, see also Figure 3.1B). Notably, this difference is most pronounced in the PUBLIC-BASE treatment, where the acquired information affects fairness perceptions and thus prosocial appearance, but monetary incentives for prosocial appearance do not confound inferences about a participant's prosocial motivation (Permutation test: \$0.36 vs. \$0.59, Z = 2.11, p = 0.033).[15] This suggests that, when the received information becomes public, especially individuals low in prosocial motivation select information for prosocial appearance.

Based on the evidence presented so far, we cannot pin down whether *low-medium* dictators are less prosocially motivated *per se*, that is, whether they make lower transfers irrespective of the transfer information they receive from their comparison dictator. Alternatively, *low-medium* dictators might also transfer less to the recipients simply because they are more likely to receive information about lower transfers, which makes their own low transfers appear more prosocial. Our information choice procedure addresses this concern directly: If dictators choose the source *low-medium*, they may with some probability receive the same information as those who choose the source *medium-medium* or the source *medium-high*, namely, the transfer information of a comparison dictator with *medium* endowment (\$2, as their own endowment). We now examine how transfers differ by information choice *conditional* on receiving transfer information of a comparison dictator with *medium* endowment, i.e., without confounds due to the endogenous selection into the information environment. Transfers do not differ between dictators choosing *medium-medium* and *medium-high* (Permutation test: \$0.61 vs. \$0.55, Z = -0.73, p = 0.471, for each treatment $p > 0.309$). Yet, in support of Hypothesis 1, dictators choosing *low-medium* transfer less than dictators choosing *medium-medium* or *medium-high* (Permutation test: mean

[14]The psychological literature associates both, 'lateral' comparisons (medium-medium) and 'upward' comparisons (medium-high), with diagnostic or 'self-assessment' motives Taylor et al. (1996).

[15]Dictators choosing *low-medium* also transfer less compared to the pool of all other dictators who choose a different a different source of information (Permutation tests: mean \$0.44 vs. \$0.57,Z = 2.57 p = 0.009). Further, they also transfer less compared to dictators choosing *medium-medium* (Permutation tests: all treatments pooled: \$0.44 vs. \$0.61, Z = 3.21, p = 0.001; PUBLIC-BASE, \$0.36 vs. \$0.59, Z = 1.99, p = 0.044) and they transfer less compared to *medium-high* dictators (Permutation tests: all treatments pooled, \$0.44 vs. \$0.56 , Z = 1.90, p = 0.055; PUBLIC-BASE, \$0.36 vs. \$0.59, Z = 1.72, p = 0.087).

$0.45 vs. $0.60, $Z = 2.14$, $p = 0.032$, see also Figure 3.2). Again, the difference between *low-medium* dictators and others is most pronounced if social information becomes public, but there are no monetary incentives to appear prosocial (PUBLIC-BASE, $0.35 vs. $0.59, Z = 1.68, $p =$ 0.097).[16]

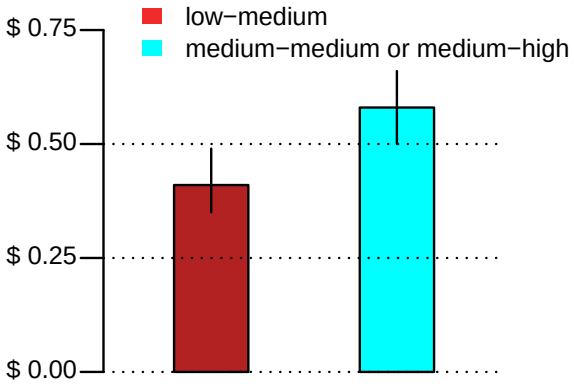


Figure 3.2: **A** Proportion of dictators that chose information from previous dictators with low endowment in Choice 1 and medium endowment in Choice 2 (low-medium), by treatment. **B** Mean transfers with bootstrapped 95% confidence intervals for each treatment, shown separately for dictators choosing information from previous dictators with low endowment in Choice 1 and medium endowment in Choice 2 (*low-medium*) or equally high or higher endowments (*medium-medium* or *medium-high*).

OLS regressions with robust standard errors yield further support for Hypothesis 1 (see Appendix Table 3.4, Models 1-6). Regardless of whether we control for the revealed transfer information, treatment dummies, or demographics, we observe that dictators choosing *low-medium* always transfer significantly less than others (Appendix 3.D, Table 3.4, $b_{low\text{-}medium}$ = -0.15, $SE = 0.06$, $p = 0.014$; also significant in Model 1-5). Result 3 sums up the evidence for Hypothesis 1.[17]

Result 3: *Dictators choosing social comparison information suggesting low prosocial behavior, i.e., low-medium, transfer less than dictators who choose information suggesting high prosocial behavior, i.e., medium-medium or medium-high. Dictators choosing information suggesting low prosocial behavior also transfer less if (by chance) they receive the same information as dictators who choose information suggesting high prosocial behavior.*

[16] We can also focus on observations where dictators receive the information that a comparison dictator with medium endowment transferred an equal split, which is a clear signal about what is fair. Dictators choosing *medium-medium* or *medium-high* do not differ in their transfer behavior ($0.67 vs. $0.78, $Z = 1.08$, $p = 0.292$), but they transfer significantly more at a 10% -level than *low-medium* dictators (Permutation test: $0.52 vs. $0.69, $Z = 1.81$, p = 0.070; PUBLIC-PUN, $0.42 vs. $0.80, $Z = 2.09$, $p = 0.039$).

[17] In line with former studies (e.g., Krupka and Weber, 2009), dictators tend to increase their transfers when the revealed transfer information is an equal split of the endowment rather than zero (Appendix 3.D, Table 3.5, Model 6, $b_{\text{Fair}} = 0.21$, SE = 0.06, $p < 0.001$, Model 1-5 also significant), regardless of the acquired information.

Lastly, looking at aggregate outcomes, our data is in line with former studies where recipients can punish dictators (for example, Fehr and Fischbacher (2004)). About 25.7% of the recipients in PRIVATE-PUN punished and, if they punished, they assigned on average 25.4 out of 100 punishment points (SD = 20.8). About 30.6% of the recipients in PUBLIC-PUN punished and, if they punished, they assigned about 27.1 points (SD = 27.3). The average punishment behavior did not differ significantly between PRIVATE-PUN and PUBLIC-PUN (6.54 vs. 8.14, Z = 0.71, p = 0.483). Transfers are higher in PRIVATE-PUN than in PRIVATE-BASE (Permutation test, \$0.48 vs. \$0.64, Z = -2.35, p = 0.019), so dictator *transfers* respond to the incentive of monetary punishment. Transfers in PUBLIC-PUN are not significantly higher than transfers in the baseline treatment PRIVATE-BASE and PUBLIC-BASE (Permutation tests, PRIVATE-BASE vs. PUBLIC-PUN, \$0.48 vs. \$0.51, Z = -0.41, p = 0.681, PUBLIC-BASE vs. PUBLIC-PUN, Permutation test, \$0.49 vs. \$0.51, Z = -0.19, p = 0.847). Transfers in PRIVATE-PUN are higher than in PUBLIC-PUN (Permutation test, \$0.64 vs. \$0.51, Z = -2.10, p = 0.036).[18] A plausible explanation for this difference is that dictators in PUBLIC-PUN can more easily excuse their low transfers when the comparison dictator has made a low transfer. This is not possible for dictators in PRIVATE-PUN, because here recipients are unaware of the dictator's social information. In line with this explanation, dictators in PUBLIC-PUN transfer less than dictators in PRIVATE-PUN, if we compare only dictators that received information about a zero transfer (Permutation test, \$0.61 vs. \$0.39, Z = 2.36, p = 0.018). We sum up the findings in Result 4.

Result 4: *In line with previous studies, recipients' punishment opportunity increases transfers, when social comparison information is private (PRIVATE-PUN). When social comparison information is public, dictators use information about comparison dictator's low transfers to excuse own low transfers, so transfers do not increase relative to the baseline (PUBLIC-PUN vs. PRIVATE-BASE).*

3.5 Discussion and Conclusion

We demonstrate that the desire to appear prosocial guides information acquisition. When the acquired information becomes public, we find that many individuals choose the source of social

[18] In both cases, the BASE and the PUN environment, selecting low-medium causally mediates the decrease in transfers (95% confidence intervals from 10,000 simulations: PRIVATE-BASE vs. PUBLIC-BASE, [-0.04, 0.00], p = 0.070; PRIVATE-BASE vs. PRIVATE-PUN, [-0.06, -0.01], p = 0.009). Yet, we hypothesize that *low-medium* dictators have less prosocial motivation per se, which introduces unobserved pre-treatment confounders and may result in biased estimates.

comparison information suggesting low prosocial behavior (i.e., *low-medium*), i.e., they prefer comparison dictators with lower endowment over comparison dictators with higher endowments and thus more likely receive a lower comparison standard. We interpret this as evidence for the motive of appearing prosocial to recipients, which is stronger under public rather than private information since the information can also affect the recipient's fairness perceptions (Battigalli and Dufwenberg, 2007; Falk et al., 2008; Andreoni and Bernheim, 2009; Heintz et al., 2015). Our results thus extend the literature on information avoidance, where the received information of the dictator always remains private information.[19] Thus, information acquisition is not only motivated by a desire to appear prosocial to oneself (Grossman and van der Weele, 2017), but also by appearing prosocial to others.

In line with the prediction of our adapted self- or social-signaling model, dictators who choose a source of information suggesting low prosocial behavior transfer significantly less than dictators interested in other sources (about 30% less than dictators choosing information suggesting high prosocial behavior). By making our information choice procedure probabilistic, subjects with different information choice motives receive the same information with some probability, i.e., transfer information from a comparison dictator with the same endowment as their own, we can compare dictator transfers conditional on receiving the same information. We find that, even when receiving the same information, people acquiring information suggesting low prosocial behavior make lower transfers than those acquiring information suggesting high prosocial behavior. We interpret this as evidence for the lower prosocial motivation of dictators acquiring information suggesting low prosocial behavior.[20]

Somewhat intriguingly, our findings address two very different perspectives on the literature on social preferences. One perspective, based on early advances in public goods studies and dictator games, suggests that heterogeneity in preferences for cooperation and giving explains the heterogeneity in social decisions. Another perspective raises concerns about whether subjects merely care about their image. This perspective is based on studies showing how prosocial

[19]Note that even in the willful ignorance setting (Dana et al., 2007), it is not possible to exclude that the desire to appear prosocial to others contributes to information avoidance since recipients know that the dictator might have been ignorant about the adverse consequences of his or her action.

[20]Importantly, our information choice procedure refines existing methods to measure heterogeneity in prosocial motivation in information acquisition. While the information avoidance literature has addressed the issue of "sorting" , the result are mixed (Larson and Capra, 2009; Fong and Oberholzer-Gee, 2011; Conrads and Irlenbusch, 2013; Grossman and van der Weele, 2017). These mixed results may have occurred because previous studies compared behavior under exogenously imposed information to behavior under information acquisition. As noted by Fong and Oberholzer-Gee (2011), this approach has methodological limitations. Researchers cannot observe an individual's counterfactual information choices under exogenous information and thus they must assume that information choices are i.i.d. across subjects and do not correlate with unobserved characteristics (e.g., prosocial motivation).

behavior in dictator games can be easily reduced by manipulating subtle cues of normative behavior (citations above and others, e.g. Bardsley (2008)). In showing how both heterogeneity in preferences and the desire to appear prosocial interact in information acquisition, the evidence presented here reconciles both perspectives.

Finally, our experiment also provides a cautious note for the design of policies involving social comparisons. Based on previous research in behavioral economics, a policy maker may deem it wise to make information about other's energy consumption, donations or paid taxes publicly available to foster socially desirable behavior (Schultz et al., 2007; Allcott, 2011). Yet, in our experiment, making social information public can also decrease overall transfers (PRIVATE-PUN vs. PUBLIC-PUN). This suggests that the proposed policy could under certain circumstances also have a detrimental effect. Individuals might selectively look to others with a high energy consumption, e.g., managers or politicians. Unfortunately, particularly these individuals are low in prosocial motivation and might take the information as evidence to excuse their low prosocial behavior, e.g., to support their belief that a high energy consumption cannot be avoided.

Appendices

3.A The Model

Our solution concept follows previous approaches by Bodner and Prelec (2003), Bénabou and Tirole (2006), Andreoni and Bernheim (2009), and, especially, (Grossman and van der Weele, 2017), who present a model of information acquisition and willful ignorance. Loosely speaking, the agent cares about how he is perceived by himself and by others, both of which he can only infer from his past actions. While the own "true" prosocial motivation v guides the own actions (i.e., the preferences are available to the "decision-maker-self"), this information is disregarded or not available when the agent (the "observer-self") reflects about his prosocial appearance (i.e., when the observer self integrates the own prosocial appearance $E(v|\ a,i)$ into the agent's utility). The observer-self evaluates the decision-maker's actions just like an outside observer, e.g., an imagined spectator or the recipient, and thus infers the agent's prosocial motivation relying only on the agent's actions. Thus, he cannot use information about the decision maker's strategy, e.g., the transfer had he acquired different information, to infer the agent's type. We assume that $\mu > 0$, so the desire to appear prosocial positively affects prosocial behavior. Further, we assume that the cost-increase due to choosing a_H over a_L is strictly larger than the maximal image gain, so $(a_H - a_L)\ c > \mu$. This excludes situations where any action could be sustained in pooling equilibriua due to image concerns alone (see Benabou and Tirole (2006)). Fairness perceptions $f(a,i) = p\ f(a - i) + (1 - p)\ f(a - a_H)$ depend on the distance of the own action to a salient norm and the social information, where $p \in (0, 0.5)$ governs the impact of the social information on fairness. Fairness is measured by f and, as Andreoni and Bernheim (2009), we assume that f is symmetric and strictly concave in $a - a_H$ with its maximum at $a = a_H$ (analogously for $f(a - i)$). The timing of the game is as follows:

1. The agent's social type v element [0,1] is drawn by nature from the distribution *g(v)* at the start of the game. *g(v)* is commonly known.

2. The agent chooses whether to acquire information source i_L or i_H, where $i_L < i_H$.

3. The agents choose whether to make a low prosocial action $a_L \geq 0$ or the normative action a_H, where $a_L < a_H$.[21]

[21] Note that we do not require that $a_L = i_L$.

4. The observer self processes the agent's transfer and the acquired information, and updates beliefs about the agent's social type. Subsequently, payoffs realize.

In what follows, we focus on separating equilibria, with a cutoff-value $v^* \in (0,1)$ such that types with a prosocial motivation $v < v^*$ acquire i_L and choose the action a_L and types with a prosocial motivation $v \geq v^*$ acquire i_H and choose the action a_H. We denote the observer-self's belief about the agent's type, i.e., the agent's image, when acquiring information i and choosing the action a with φ_a^i. Beliefs on the equilibrium path are $\varphi_{a_L}^{i_L} = \int_0^{v^*} vg(v)\mathrm{d}v/G(v^*)$ and $\varphi_{a_H}^{i_H} = \int_{v^*}^1 vg(v)\mathrm{d}v/(1-G(v^*))$, so $\varphi_{a_H}^{i_H} > \varphi_{a_L}^{i_L}$ for any v^*. Off the equilibrium path, we assume that beliefs are weakly increasing in a. Further, $\varphi_{a_L}^{i_L} \geq \varphi_{a_L}^{i_H}$ and $\varphi_{a_H}^{i_H} \geq \varphi_{a_H}^{i_L}$. We later show that these beliefs are reasonable using the pD1 refinement as defined in, e.g., Grossman and van der Weele (2017) and Andreoni and Bernheim (2009).

Now we establish conditions such that, given their chosen information, agents do not deviate from the equilibrium path. Given that an agent has chosen i_L, he weakly prefers a_L over a_H if $v \leq ((a_H - a_L)c - \mu(\varphi_{a_H}^{i_L} - \varphi_{a_L}^{i_L}))/(\mathrm{f}(a_H, i_L) - f(a_L, i_L)) \equiv v^L$, where $v^L > 0$, since we assumed $(a_H - a_L)\ \mathrm{c} > \mu$. Given that an agent has chosen i_H, he weakly prefers a_H over a_L if $v \geq ((a_H - a_L)c - \mu(\varphi_{a_H}^{i_H} - \varphi_{a_L}^{i_H}))/(\mathrm{f}(a_H, i_H) - f(a_H, i_L)) \equiv v^H$.

Agents above or at the threshold v^* (who choose (i_H, a_H)) do not deviate to choosing (a_L, i_H), if $v^* > v^H$. Since f(a_H,i_H) > f(a_L,i_L) > f(a_L,i_H), this holds true since we assumed that $\varphi_{a_L}^{i_L} \geq \varphi_{i_H}^{a_L}$. As a result, the agents with lower prosocial motivation that would not prefer a_H over a_L given i_H rather choose i_L than i_H in the first place.

Agents below the threshold v^* (who choose (i_L, a_L)) do not deviate to choosing a_H, if $v^* > v^{\mathrm{L}}$. The inequality $v^* > v^{\mathrm{L}}$ holds true since f(a_H,i_H) > f(a_H,i_L) > f(a_L,i_L) and we assumed that $\varphi_{a_H}^{i_H} \geq \varphi_{a_H}^{i_L}$. As a result, the agents with higher prosocial motivation that would not prefer a_L over a_H given i_L rather choose i_H than i_L in the first place.

In equilibrium, the more prosocially motivated agents should choose the information suggesting high transfers and make high transfers (i_H, a_H), whereas the less prosocially motivated agents should choose information suggesting low transfers and make low transfers (i_L, a_L). For agents to weakly prefer (i_H, a_H) over (i_L, a_L), we need that the expected utility of choosing (i_H, a_H) is greater or equal than the expected utility of choosing (i_L, a_L). This is the case if $v \geq v^*$, where v^* is defined implicitly by the fixed point equation $v^* = ((a_H - a_L)c - \mu(\varphi_{a_H}^{i_H} - \varphi_{a_L}^{i_L}))/(\mathrm{f}(a_H, i_H) - f(a_L, i_L))$. Existence follows from continuity of both sides of this equation. To ease exposition, we assume that types at the threshold who are indifferent choose (i_H, a_H).

We next show sufficient conditions for a threshold $v^* \in (0,1)$ to exist. The threshold v^* is bounded above 0 for any beliefs because we assumed $\mu < (a_H - a_L)$ c $\equiv \overline{\mu}$. Further, v^* is bounded below 1, if $\mu >$ (($a_H - a_L$) c - (f(a_H , i_H) - f(a_L, i_L)))/(1 – E(v)) $\equiv \underline{\mu}$, where we use that given $v^* = 1$, $\varphi_{a_H}^{i_H} = 1$ and $\varphi_{a_H}^{i_H} =$ E(v). If we assume c ≤ 1, a sufficient condition for $\underline{\mu} < \mu < \overline{\mu}$ to be true, so $v^* \in (0,1)$ exists, is that ($a_H - a_L$) c < f(a_H, i_H) – f(a_L, i_L), which holds since we assumed strict concavity and symmetry of f w.r.t. $a - a_H$ and $a - i$, respectively, and $p < 1/2$ and $\mu > 0$. Otherwise if c > 1, a sufficient condition for $\underline{\mu} < \mu < \overline{\mu}$, so $v^* \in (0,1)$ exists, is that (f(a_H, i_H) – f(a_L, i_L))/(E(v) ($a_H - a_L$)) > c.

Taken together, the above arguments have shown that an interior separating equilibrium exists if $\underline{\mu} < \mu < \overline{\mu}$ and $p< 1/2$. In any such equilibrium the type at the cutoff v^* must be indifferent between choosing (a_H,i_H) over (a_L,i_L), i.e., the payoff difference z between choosing (a_H,i_H) over (a_L,i_L) must equal zero. A sufficient condition to ensure the stability (uniqueness) of the threshold is that the the total derivative dz of the payoff difference z, with respect to v^* be positive (positive for all v). This is the case if the image difference between agents choosing (a_H,i_H) and agents choosing (a_L,i_L) does not decrease too strongly in v^*, i.e., if d($\varphi_{a_H}^{i_H} - \varphi_{a_L}^{i_L}$)/d$v^*$ > - (f(a_H,i_H) - f(a_L,i_L))/μ. For this to hold we need that the distribution of social types $g(v)$ is not increasing too steeply anywhere on its domain, e.g., agents with a high type (prosocial motivation) should not be too frequent relative to agents with a low type (prosocial motivation); see Benabou and Tirole (2006), pp. 1666-1667, for a thorough discussion of the implications of different distributions.

Next, we demonstrate the reasonableness of our assumed off-equilibrium beliefs. We use the pD1 refinement (e.g., as defined by Grossman and van der Weele (2017), p. 182). To pin down off-equilibrium beliefs, we need to find out which of our types deviate for the largest set of off-equilibrium beliefs. To this end, we first establish conditions such that types above or at the threshold deviate from their equilibrium behavior and then establish conditions such that types below the threshold deviate from their equilibrium behavior. Then, we compare the resulting expressions to see which of our types deviate for a larger set of off-equilibrium beliefs.

We start with $\varphi_{a_L}^{i_L} \geq \varphi_{a_L}^{i_H}$. Agents above or at the threshold deviate if their equilibrium payoffs are lower than their deviation payoffs, that is, if $\mathrm{f}(a_H, i_H)v + \mu\varphi_{a_H}^{i_H} - a_H c < \mathrm{f}(a_L, i_H)v + \mu\varphi_{a_L}^{i_H} - a_L c$. Hence, they deviate if $\varphi_{a_L}^{i_H} > \varphi_{a_H}^{i_H} + [v(\mathrm{f}(a_H, i_H) - \mathrm{f}(a_L, i_H)) - (a_H - a_L)c]/\mu \equiv \varphi_{a_L, i_H}^{v \geq v^*}$. Agents below the threshold deviate if their equilibrium payoffs are lower than their deviation payoffs, that is, if $\mathrm{f}(a_L, i_L)v + \mu\varphi_{a_L}^{i_L} - a_L c < \mathrm{f}(a_L, i_H)v + \varphi_{a_L}^{i_H} - a_L c$. Hence, they deviate if $\varphi_{a_L}^{i_H} > \varphi_{a_L}^{i_L} + [v(\mathrm{f}(a_L, i_L) - \mathrm{f}(a_L, i_H))]/\mu \equiv \varphi_{a_L, i_H}^{v < v^*}$. Comparing the two expressions, we see that

$\varphi_{a_L,i_H}^{v\geq v^*} \geq \varphi_{a_L,i_H}^{v<v^*}$, if $v \geq ((a_H - a_L)c - \mu(\varphi_{a_H}^{i_H} - \varphi_{a_L}^{i_L}))/(\mathrm{f}(a_H, i_H) - f(a_L, i_L))$, where the right-hand side is the same expression as v^* and thus satisfied in equilibrium. This means that the lower types deviate for a larger set of off-equilibrium beliefs, so any belief below or equal to v^* is reasonable and $\varphi_{a_L}^{i_H} = 0$ is the lowest belief that can be supported by the pD1 refinement.

Next, we consider $\varphi_{a_H}^{i_H} \geq \varphi_{a_H}^{i_L}$. Agents above or at the threshold deviate if their equilibrium payoffs are lower than their deviation payoffs, that is, if $\mathrm{f}(a_H, i_H)v + \mu\varphi_{a_H}^{i_H} - a_H c < \mathrm{f}(a_H, i_L)v + \mu\varphi_{a_H}^{i_L} - a_H c$. Hence, they deviate if $\varphi_{a_H}^{i_L} > \varphi_{a_H}^{i_H} + [v(\mathrm{f}(a_H, i_H) - \mathrm{f}(a_H, i_L))]/\mu \equiv \varphi_{a_H,i_L}^{v\geq v^*}$. Agents below the threshold deviate if their equilibrium payoffs are lower than their deviation payoffs, that is, if $\mathrm{f}(a_L, i_L)v + \mu\varphi_{a_L}^{i_L} - a_L c < \mathrm{f}(a_L, i_H)v + \mu\varphi_{a_L}^{i_H} - a_L c$. Hence, they deviate if $\varphi_{a_H}^{i_L} > \varphi_{a_L}^{i_L} + [v(\mathrm{f}(a_L, i_L) - \mathrm{f}(a_H, i_L)) - (a_H - a_L)c]/\mu \equiv \varphi_{a_H,i_L}^{v<v^*}$. Comparing the two expressions, we see that $\varphi_{a_H,i_L}^{v\geq v^*} \geq \varphi_{a_H,i_L}^{v<v^*}$, if $v \geq ((a_H - a_L)c - \mu(\varphi_{a_H}^{i_H} - \varphi_{a_L}^{i_L}))/(\mathrm{f}(a_H, i_H) - f(a_L, i_L))$, where the right-hand side is the same expression as v^* and thus satisfied in equilibrium. This means that the lower types deviate for larger set of off-equilibrium beliefs, so any belief below or equal to v^* is reasonable and $\varphi_{a_H}^{i_L} = 0$ is the lowest belief that can be supported by the pD1 refinement.

Note that our separating equilibrium implies Hypothesis 1: Agents acquiring information i_L have a lower prosocial motivation and make lower transfers than agents acquiring i_H. Would the two types both nevertheless receive i_H, those choosing i_L would nevertheless choose the lower transfer. Their choice of i_L merely helps them to improve their fairness perceptions $\mathrm{f}(a_L,i_L)$ and prosocial appearance $\varphi_{a_L}^{i_L}$.

We now turn to the comparative statics which are important to derive Hypotheses 2 and 3.[22] Hypothesis 2 implies that if we make social information public and thus increase the impact of the information on fairness perceptions p, the threshold v^* should increase, i.e., more agents should be willing to acquire information suggesting low transfers and make a lower transfer. Under the assumption that the equilibrium is stable, we analyze the change in v^* caused by a change in p by using the total derivative of the payoff difference between choosing $(a_H\,,\, i_H)$ and choosing $(a_L,\, i_L)$ with respect to p and set it to zero, i.e., $\mathrm{d}z_p{=}0$, and solve for $\mathrm{d}v^*\,/\mathrm{dp}$. The resulting expression $\mathrm{d}v^*\,/\mathrm{dp} = ((\mathrm{f}(a_H\text{-}a_H) - \mathrm{f}(a_L\text{-}a_H))\; v^*\;)/(\mathrm{f}(a_H,i_H) - \mathrm{f}(a_L,i_L) + \mathrm{d}(\varphi_{a_H}^{i_H} - \varphi_{a_L}^{i_L})/\mathrm{d}v^*)$ is larger than 0 if and only if $\mathrm{d}(\varphi_{a_H}^{i_H} - \varphi_{a_L}^{i_L})/\mathrm{d}v^* > -(\mathrm{f}(a_H,i_H) - \mathrm{f}(a_L,i_L))/\mu$. This is the same condition that is required for uniqueness and stability of the threshold and thus satisfied in equilibrium. Hypothesis 3 also implies that if we introduce punishment and thus increase the payoff-related image motivation, then the threshold v^* should decrease. We set the total derivative with

[22] It is also possible to extend our setup to a situation where, with some probability q, information source i_L yields the same information as information source i_H, but different information with probability 1-q, without affecting the comparative statics.

respect to μ equal to zero, dz_μ=0 and solve for dv^* /$d\mu$. The resulting expression dv^* /$d\mu$ = - $(\varphi^{i_H}_{a_H} - \varphi^{i_L}_{a_L})$/(f($a_H$,$i_H$) - f($a_L$,$i_L$) + μ d$(\varphi^{i_H}_{a_H} - \varphi^{i_L}_{a_L})/dv^*$) is smaller than 0 if and only if $d(\varphi^{i_H}_{a_H} - \varphi^{i_L}_{a_L})/dv^* > -$ (f(a_H,i_H) - f(a_L,i_L))/μ. Again, this is the same condition that is required for uniqueness and stability of the threshold and thus satisfied in equilibrium.

3.B Demographic Controls

	Age	Male	Experience
PRIVATE-BASE	33.16 (11.05)	0.59 (0.49)	4.46 (13.62)
PRIVATE-PUN	34.11 (10.94)	0.49 (0.50)	2.86 (10.39)
PUBLIC-BASE	32.91 (9.20)	0.54 (0.50)	1.71 (2.79)
PUBLIC-PUN	32.50 (8.69)	0.56 (0.50)	1.42 (10.03)
K-Sample Perm. Test	0.704	0.557	0.020

Table 3.2: Demographic controls Age is the self-reported age in years; Male indicates whether the participant self-reports to be male; Experience is the self-reported number of participations in similar games. In one of three cases, covariates are not balanced. Notably, including this covariate does not change significance patterns when included in analyses.

3.C Chosen source of social comparison information

Chosen sources	low-medium		medium-medium		medium-high	
	(1)	(2)	(3)	(4)	(5)	(6)
Public	0.5241°	0.6447*	−0.0869	−0.1080	−0.4115	−0.5143
	(0.3041)	(0.3196)	(0.2836)	(0.2923)	(0.3344)	(0.3413)
Pun	−0.3378	−0.2480	0.1080	0.0436	0.2432	0.1384
	(0.3366)	(0.3451)	(0.2828)	(0.2919)	(0.3088)	(0.3252)
Public*Pun	0.4364	0.3636	−0.7161°	−0.6368	0.1582	0.2054
	(0.4438)	(0.4531)	(0.4144)	(0.4231)	(0.4571)	(0.4679)
Age		0.0104		0.0220*		−0.0280*
		(0.0110)		(0.0104)		(0.0123)
Male		0.3410		−0.3761°		−0.0798
		(0.2276)		(0.2113)		(0.2337)
Exper		0.0222°		0.0064		−0.0725*
		(0.0127)		(0.0134)		(0.0343)
(Intercept)	0.5001***	0.2564**	0.5733***	0.2756*	0.4648***	0.2123*
	(0.0469)	(0.0930)	(0.0489)	(0.1089)	(0.0637)	(0.1063)
Treatment effects						
PUBLIC-BASE vs.	0.5241°	0.6447*	-0.0869	-0.1080	-0.4115	-0.5143
PRIVATE-BASE	(0.3041)	(0.3196)	(0.2836)	(0.2923)	(0.3344)	(0.3414)
PUBLIC-PUN	0.9605***	1.0083***	-0.8030***	-0.7448**	-0.2533	-0.3088
vs. PRIVATE-PUN	(0.3217)	(0.3257)	(0.3007)	(0.3071)	(0.3100)	(0.3167)
Log Likelihood	-246.2405	-240.5809	-268.9922	-262.6716	-233.6081	-226.0425
Num. obs.	411	408	411	408	411	408

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\circ}p < 0.1$

Table 3.3: Binomial logit regressions with intercept and robust standard errors predicting the probability that an individual chooses the sources of social comparison indicated in the column header as opposed to any other available social comparison. Estimates are based on models with treatment dummies Public, Pun, Public*Pun as predictors and, if included, with demographic controls Age, Male, Experience. The Treatment effect BASE (PUBLIC – PRIVATE) tests the null hypothesis that the coefficient Public=0. The Treatment effect PUN (PUBLIC – PRIVATE) tests the null hypothesis that the coefficients Public + Public*Pun=0. Pun indicates whether recipients could punish dictators, Public indicates whether recipients were informed about comparison allocations. Regressions (2), (4), and (6) exclude three observations from individuals who did not respond to the demographic questions in models with demographic controls.

3.D Transfers

	Transfers					
	(1)	(2)	(3)	(4)	(5)	(6)
Fair	0.1840**	0.1976***			0.1849**	0.1975***
	(0.0567)	(0.0551)			(0.0566)	(0.0549)
Low-medium	−0.1410*	−0.1502*	−0.1399*	−0.1557*	−0.1319*	−0.1457*
	(0.0603)	(0.0597)	(0.0627)	(0.0621)	(0.0616)	(0.0613)
Age		0.0082***		0.0088***		0.0081**
		(0.0024)		(0.0025)		(0.0025)
Male		−0.0866		−0.0502		−0.0796
		(0.0546)		(0.0570)		(0.0545)
Experience		0.0016		0.0028		0.0022
		(0.0033)		(0.0034)		(0.0034)
Public			−0.0307	0.0167	−0.0060	0.0223
			(0.0589)	(0.0817)	(0.0809)	(0.0793)
Pun			0.0713	0.0928	0.0966	0.0939
			(0.0568)	(0.0754)	(0.0771)	(0.0743)
Public*Pun				−0.0610	−0.0457	−0.0647
				(0.1123)	(0.1113)	(0.1095)
(Intercept)	0.5001***	0.2564**	0.5733***	0.2756*	0.4648***	0.2123*
	(0.0469)	(0.0930)	(0.0489)	(0.1089)	(0.0637)	(0.1063)
R^2	0.0537	0.1042	0.0237	0.0685	0.0613	0.1096
Adj. R^2	0.0467	0.0872	0.0128	0.0435	0.0437	0.0822
Num. obs.	272	269	272	269	272	269

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\circ}p < 0.1$

Table 3.4: Predicting transfers when dictators receive transfer information of a comparison dictator who has the same endowment ($2). OLS regressions with intercept and robust standard errors predicting dictator transfers when the received transfer information for comparison is from a comparison dictator who has the same endowment, i.e., *medium* ($2). Receiving information about a transfer of zero in PRIVATE-BASE serves as the reference category. Fair indicates whether the comparison dictator transferred the equal split of that dictators endowment. Low-medium indicates that the dictator chose the information source *low-medium*. Estimates are based on models with treatment dummies Public, Pun, Public*Pun and Low-medium as predictors and, if included, the demographic controls Age, Male, Experience. Pun indicates whether recipients could punish dictators, Public indicates whether recipients were informed about the comparison dictator's transfer. Regressions (2), (4), and (6) exclude three observations from individuals who did not respond to demographic questions.

Transfers	low-medium		medium-medium		medium-high	
	(1)	(2)	(3)	(4)	(5)	(6)
Low Fair	0.1485	0.1786*				
	(0.0935)	(0.0854)				
Medium Zero	0.0379	0.0190				
	(0.0944)	(0.0894)				
Medium Fair	0.1666°	0.1187	0.1339°	0.1792*	0.4104**	0.4054**
	(0.0972)	(0.0995)	(0.0774)	(0.0753)	(0.1525)	(0.1522)
High Zero					0.2059	0.2091
					(0.1472)	(0.1477)
High Fair					0.1557	0.1567
					(0.1367)	(0.1360)
Age		0.0135***		0.0063*		0.0051
		(0.0032)		(0.0031)		(0.0062)
Male		0.0427		−0.1305°		−0.0680
		(0.0671)		(0.0736)		(0.1066)
Experience		0.0014		0.0041		−0.0035
		(0.0029)		(0.0035)		(0.0135)
Public	−0.0233	0.0337	0.0487	0.0661	0.0074	0.0179
	(0.1015)	(0.0941)	(0.1109)	(0.1081)	(0.1605)	(0.1668)
Pun	0.1986	0.1710	0.1426	0.1331	0.0226	0.0090
	(0.1218)	(0.1170)	(0.0935)	(0.0905)	(0.1525)	(0.1567)
Public*Pun	−0.0926	−0.1036	−0.0517	−0.0528	−0.1536	−0.1445
	(0.1479)	(0.1427)	(0.1501)	(0.1492)	(0.2118)	(0.2137)
(Intercept)	0.3006**	−0.1978	0.4582***	0.2682°	0.4052**	0.2852
	(0.0937)	(0.1337)	(0.0822)	(0.1369)	(0.1319)	(0.2423)
R^2	0.0750	0.2028	0.0347	0.0846	0.0744	0.0863
Adj. R^2	0.0279	0.1393	0.0093	0.0413	0.0189	0.0015
Num. obs.	125	123	157	156	107	107

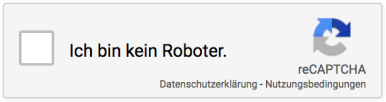
$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\circ}p < 0.1$

Table 3.5: Predicting transfers by chosen source of social comparison information. OLS regressions with intercept and robust standard errors predicting dictator transfers. We ran separate regressions for each subset of dictators choosing the source of information comparison indicated in the column header. Medium Fair (Medium Zero) indicates whether the dictator received information about an equal split transfer (a zero transfer) of a comparison dictator with a medium endowment ($2). The names of the other predictors are analogous. Low Zero serves as the reference observation for subjects choosing low-medium, Medium Zero as the reference observation for medium-medium and medium-high models. Estimates are based on models with treatment dummies Public, Pun, Public*Pun as predictors and, if included, the demographic controls Age, Male, Experience. Pun indicates whether recipients could punish dictators, Public indicates whether the information was made public. Regression (2) excludes two observations and regression (4) excludes one observation from individuals who did not respond to demographic questions in models with demographic controls.

3.E Instructions

All treatments

Please solve the recognition task below.



Please enter your Worker ID:

ID

Your Worker ID is needed to ensure you get your payoff.
If you don't know your Worker ID you can find it out on the following page (opens in a new tab):
https://www.mturk.com/mturk/dashboard

Next

Anonymity

Your MTurk Worker ID and total payoffs are confidential: only you and the creator of the survey will know the amount of money you make.

Back Next

Payment

You will participate in an online game. Your *participation payoff* is $1.

In addition, you can earn a *payoff from playing the game*. The *payoff from playing the game* will be granted as a **bonus payment within 24h** after the HIT is completed.

Your total payoff = participation payoff + payoff from playing the game

Back Next

You will be randomly assigned the role of Player A or Player B.

Back Next

Instructions Player A

Please read closely.

You are Player A. You will be allocated a Bonus of either $1, $2 or $4. You can choose how much of this bonus to TRANSFER to Player B. Player B is also an MTurk Worker.

Payoff Player A = Bonus – TRANSFER to Player B
Payoff Player B = TRANSFER by Player A

Player B WILL NOT know the actual Bonus allocated to you ($1, $2 or $4).
BUT Player A and Player B WILL see this identical instruction screen.

Next

Question

Please answer the following question correctly to ensure your HIT is accepted. You have two trials to select the correct answer. To answer, please choose one option.

You will choose how much to TRANSFER to Player B. What is Player B's information on the size of your Bonus?

- ○ Player B will know my Bonus is either $1, $2, or $4
- ○ Player B will know nothing about my Bonus
- ○ Player B will know the size of my Bonus

Back Next

Bonus Allocation

The Bonus allocated to you is:

$2

Later, you will choose how much of this Bonus to TRANSFER to Player B. Player B WILL NOT know the size of your Bonus.

Back Next

PUBLIC-BASE and PUBLIC-PUN

Information about a PREVIOUS GAME

Before making a TRANSFER, you will see INFORMATION on the Transfer from a Player A to a Player B on a INFORMATION DISPLAY.

The INFORMATION will be randomly selected from a PREVIOUS GAME.

The INFORMATION DISPLAY will look like this:

A previous Player A has transferred to a previous Player B:

$ [INFORMATION]

Player B WILL also see the INFORMATION DISPLAY.
Player A and Player B WILL see this identical instruction screen.

Back Next

PRIVATE-BASE and PRIVATE-PUN

Information about a PREVIOUS GAME

Before making a TRANSFER, you will see INFORMATION on the Transfer from a Player A to a Player B on the INFORMATION DISPLAY.

The INFORMATION will be randomly selected from a PREVIOUS GAME.

The INFORMATION DISPLAY will look like this:

A previous Player A has transferred to a previous Player B:

$ [INFORMATION]

Player B WILL NOT see the INFORMATION DISPLAY.
Player B WILL NOT see this instruction screen.

Back Next

PUBLIC-PUN

Instructions Player A

Please read closely.

Below you see the instructions of Player B.

Game Instructions Player B

Deduction by Player B

You will see the TRANSFER by Player A and the INFORMATION DISPLAY.

Afterwards, you, but also Player A, will receive an Endowment of $1. Then, ONLY you can use the Endowment to assign DEDUCTION Cents to Player A.

Each DEDUCTION Cent will cost you $0.01 and reduce the payoff of Player A by $0.03

Payoff Player A = Bonus – TRANSFER to Player B
*+ Endowment - 3*DEDUCTION Cents*

Payoff Player B = TRANSFER by Player A
+ Endowment - DEDUCTION Cents

Back Next

Question

Please answer the following question correctly to ensure your HIT is accepted. A correct answer may require selecting more than one option. You have three trials to select the correct option(s).

The following statements are about the DEDUCTION. Which of them are correct?

- [ ] If Player B assigns 1 DEDUCTION Cent to Player A, this reduces the Payoff of Player A by $0.01.
- [ ] Player B will first see the INFORMATION DISPLAY and TRANSFER. Afterwards, Player B can assign DEDUCTION Cents to Player A.
- [ ] Player A can also assign DEDUCTION Cents to Player B.
- [ ] Player A and B will receive an Endowment of $1.

Back Next

PRIVATE-PUN

Instructions Player A

Please read closely.

Below you see the instructions of Player B.
Note that Player B will not see the INFORMATION DISPLAY, but only your TRANSFER.

Game Instructions Player B

Deduction by Player B

You will see the TRANSFER by Player A.

Afterwards, you, but also Player A, will receive an Endowment of $1. Then, ONLY you can use the Endowment to assign DEDUCTION Cents to Player A.

Each DEDUCTION Cent will cost you $0.01 and reduce the payoff of Player A by $0.03

Payoff Player A = Bonus – TRANSFER to Player B
*+ Endowment - 3*DEDUCTION Cents*

Payoff Player B = TRANSFER by Player A
+ Endowment - DEDUCTION Cents

Back Next

Question

A correct answer may require selecting more than one option! Please answer the following question correctly to ensure your HIT is accepted. You have four trials to select the correct option(s).

The following statements are about the DEDUCTION. Which of them are correct?

- [ ] Player A and B will receive an Endowment of $1.
- [ ] Player B will not see the INFORMATION DISPLAY.
- [ ] If Player B assigns 1 DEDUCTION Cent to Player A, this reduces the Payoff of Player A by $0.01.
- [ ] Player B will first see the TRANSFER. Afterwards, Player B can assign DEDUCTION Cents to Player A.
- [ ] Player A can also assign DEDUCTION Cents to Player B.

Back Next

All treatments

Information about a PREVIOUS GAME

Please read closely.

You can select the INFORMATION about the Transfer of a previous Player A shown on the INFORMATION DISPLAY. The INFORMATION differs only in terms of the Bonus of the previous Player A.

You can make two selections from each of the following information menus.

Information Menu 1: INFORMATION on the Transfer of
Player A with Bonus $1
Player A with Bonus $2

Information Menu 2: INFORMATION on the Transfer of
Player A with Bonus $2
Player A with Bonus $4

One of these two information menus will be randomly chosen for the INFORMATION DISPLAY.

Back Next

PRIVATE-BASE and PRIVATE-PUN

Please note that:
Player B WILL NOT know that you selected the INFORMATION on the INFORMATION DISPLAY and Player B WILL NOT know about the INFORMATION DISPLAY.

Recall that the INFORMATION DISPLAY will look like this:

A previous Player A has transferred to a previous Player B:

$ [INFORMATION]

Note that the INFORMATION DISPLAY does not show the Bonus of a Player A from a PREVIOUS GAME.

Back Next

Question

A correct answer may require selecting more than one option! Please answer the following question correctly to ensure your HIT is accepted. You have three trials to select the correct option(s).
The following statements are about the information from a PREVIOUS GAME. Which of them are correct?

- [ ] Player B will also see the INFORMATION DISPLAY.
- [ ] The INFORMATION DISPLAY will show the Transfer of a Player A from a PREVIOUS GAME.
- [ ] Player B WILL know that you selected the INFORMATION on the INFORMATION DISPLAY.
- [ ] The INFORMATION DISPLAY will not show the Bonus of a Player A from a PREVIOUS GAME.

Back Next

PUBLIC-BASE and PUBLIC-PUN

Please note that:
Player B WILL NOT know that you selected the INFORMATION on the INFORMATION DISPLAY, but Player B WILL see the INFORMATION DISPLAY.

Recall that the INFORMATION DISPLAY will look like this:

A previous Player A has transferred to a previous Player B:

$ [INFORMATION]

Note that the INFORMATION DISPLAY does not show the Bonus of a Player A from a PREVIOUS GAME.

Back Next

Question

Please answer the following question correctly to ensure your HIT is accepted. A correct answer may require selecting more than one option. You have three trials to select the correct option(s).
The following statements are about the information from a PREVIOUS GAME. Which of them are correct?

- [ ] Player B WILL know that you selected the INFORMATION on the INFORMATION DISPLAY.
- [ ] The INFORMATION DISPLAY will show the Transfer of a Player A from a PREVIOUS GAME.
- [ ] Player B will also see the INFORMATION DISPLAY.
- [ ] The INFORMATION DISPLAY will not show the Bonus of a Player A from a PREVIOUS GAME.

Back Next

All treatments

Information about a PREVIOUS GAME

Please select the INFORMATION on the Transfer of a Player A you would like to see on the INFORMATION DISPLAY.

Information Menu 1: INFORMATION on the Transfer of

Player A with Bonus $2
Player A with Bonus $1

Information Menu 2: INFORMATION on the Transfer of

Player A with Bonus $2
Player A with Bonus $4

Back Next

PUBLIC-BASE and PUBLIC-PUN

Information Menu 1 was randomly chosen. Player B will also see the INFORMATION DISPLAY.

INFORMATION DISPLAY:

A previous Player A has transferred to a previous Player B:

$0.50

Please note:
Player B does not know that the Player A from the PREVIOUS GAME had a Bonus of $1.

You have a bonus of $2.

Please indicate the amount ($x.xx) you want to TRANSFER to Player B.

Submit

PRIVATE-BASE and PRIVATE-PUN

Information Menu 1 was randomly chosen. Player B WILL NOT see the INFORMATION DISPLAY.

INFORMATION DISPLAY:

A previous Player A has transferred to a previous Player B:

$0.00

Please note:
Player A from the PREVIOUS GAME had a Bonus of $1.

You have a bonus of $2.

Please indicate the amount ($x.xx) you want to TRANSFER to Player B.

Submit

PRIVATE-PUN and PUBLIC-PUN (adapted in BASE treatments)

Total Payoff

Your *payoff from playing the game* is $4. This is your bonus payment.
Your participation payoff is $1.

You will find out about your assigned DEDUCTION Cents by Player B in the message accompanying your bonus payment.

*Your total payment = $5 - 3*DEDUCTION Cents by Player B*

Demographic Questions

In how many transfer games like this have you participated before?

How old are you (in years)?

Please indicate your sex. (f/m/other)

Next

Chapter 4

Social-comparison engineering in children

4.1 Introduction

Performance evaluations are ubiquitous in today's social and economic life. Admissions, promotions, grades, as well as the allocation of various benefits regularly rely on relative evaluations comparing several candidates. An individual's performance depends on the rules of the evaluation procedure and the other persons he or she is compared with, i.e., the comparison standard. Indeed, a large body of evidence shows that the comparison standard strongly affects evaluators' perceptions at school (Marsh and Parker, 1984; Marsh et al., 2000), in university (Marsh et al., 2000), or at work (Garcia et al., 2010; Clark and Senik, 2010; Babcock et al., 1996; Prendergast and Topel, 1993). At school, for example, pupils with low- rather than high-performing peers not only view their own academic performance more positively, but, for a given level of performance, they are also more likely to receive positive evaluations (Marsh et al., 2000). Alike, receiving extra care or social benefits may be more likely when others appear to perform better than oneself.

To receive positive relative evaluations, people may lobby for changes in the rules of the evaluation procedure (e.g. Abbink and Hennig-Schmidt, 2006) or engage in outright sabotage (e.g. Carpenter et al., 2010; Harbring and Irlenbusch, 2011). Alternatively, a strategic selection of favorable comparison standards, i.e., social-comparison engineering, may be a possible remedy. By selecting a lower comparison standard, an individual may not only improve his or her chances of receiving favorable relative evaluations, but, most likely, may also improve the own social image

(Leary and Baumeister, 2000; Rudolph et al., 2005; Frank, 1985) and personal (economic) well-being (Marsh et al., 2000; Crocker and Canevello, 2008; Verduyn et al., 2015; Frank, 1985; Clark and Senik, 2010). For instance, to increase his or her chances of receiving positive evaluations, a researcher may select a good rather than an excellent faculty to receive higher approval from the other faculty members and potentially higher chances of promotion. Similarly, pupils or students may choose which class or sports team to join, or, simply, which classmates to sit next to.

Yet, social-comparison engineering may also come with the downside of creating an internal conflict between the individual's preference for a favorable evaluation and the preference for certain strongly favored comparison standards: For example, researchers may strive to be part of a highly-respected faculty, and pupils may desire to sit next to 'the better ones'.

While adults have been found to seek (avoid) situations that will likely lead to favorable (unfavorable) evaluations (Garcia et al., 2013; Tesser, 1988), research on children's behavior in such situations is missing yet. Previous research on children's behavior solely focused on how children perform in competitive situations (e.g., Gneezy and Rustichini, 2004; Almås et al., 2015) or how they conform to social expectations, e.g., to manage their reputation (Banerjee, 2000; Engelmann et al., 2012). Yet, up to now, it is unknown whether and from which age on children actively shape their comparison standards in relative evaluations, i.e., situations where others make decisions relevant to them. Also, little is known about the motivational and cognitive processes supporting children's ability to shape their comparison standards.[1] In this paper, we therefore investigate experimentally when and how the ability of social-comparison engineering develops.

To study systematically the development of social-comparison engineering, we let children from 4 to 9 years of age select their comparison standards in an economic experiment. To create a performance evaluation of high relevance to children, we employed children's drawings (Butler, 1989; Fu and Lee, 2007; Chevallier et al., 2012). First, each child made three drawings from three sample photographs (e.g., tree, house, and bike). Then, for each of the child's drawings, each child privately chose one out of the two other drawings to be positioned side by side to the child's own drawing. By design, one other drawing was of low accuracy and the other one was of high accuracy, such that the accuracy of the child's drawing of the photograph was in between. To elicit children's preferences for the other drawings, in the *Baseline* treatment, the

[1]There is a large body of developmental research investigating children's allocations of valuable resources within social interactions (Sutter et al., 2019). This research suggests that it is rather the underlying motivations (e.g., selfishness, egalitarianism) than cognitive abilities that cause substantial changes in allocation behavior during childhood (Almås et al., 2010; Blake et al., 2014; Buttelmann and Böhm, 2014; Fehr et al., 2008).

child's choice had no material consequences. In the two other treatments, the child's drawing was evaluated relative to the other drawing and this evaluation determined whether or not the child won a prize.

In the treatment *BetterWins*, children learned that they would win a prize if the own drawing was more accurate than the other drawing. Thus, in order to win the prize, the child had to choose the less accurate other drawing as the comparison standard. In contrast, in the treatment *WorseWins*, children learned that they would win a prize if the own drawing was less accurate than the other drawing. Thus, in order to win the prize, the child had to choose the more accurate drawing as the comparison standard. The *Baseline* informs us about children's basic preference for either high- or low-accuracy other drawings. When children in *Baseline* reveal a preference for the high-accuracy other drawing, *BetterWins* generates an internal conflict (because winning the prize required selecting low-accuracy rather than high-accuracy other drawings), while *WorseWins* does not. Yet, when children in *Baseline* reveal a preference for the low-accuracy other drawing, *WorseWins* generates an internal conflict (because winning the prize required selecting high-accuracy rather than low-accuracy other drawings), while *BetterWins* does not. Notably, to pin down the extent to which motivational and cognitive processes explain children's choices, we also elicited children's motivation to win, their desire to be liked by others, and a measure of self-control (inhibitory control in the day-night stroop task (Gerstadt et al., 1994)).

We find that in *Baseline* children strongly prefer the high-accuracy other drawing, regardless of their age. Notably, most children refer to the other drawing's beauty when explaining their choice. This suggests that children face an internal conflict in *BetterWins*, but not in *WorseWins*. Strikingly, in *BetterWins* only children from about 6-7 years of age win the prize in about 79% of the trials whereas younger children's behavior does not differ from random choice (49% of the trials). In *WorseWins*, by contrast, older and younger children win similarly often and they win in about 75% of the trials. Further analyses suggest that the age differences found in *BetterWins* can best be attributed to children's improved self-control, which might have enabled them to overcome their preference for high-accuracy other drawings. We thus present first experimental evidence that self-control is not only important for individual decision-making, but also for (the development of) strategic behavior in performance evaluations and, perhaps, social situations more generally.

4.2 Experiment

4.2.1 Participants and Procedure

Experimental sessions took place in 7 kindergartens and 6 schools, with a maximum of two sessions per institution. The sample consisted of N = 172 children. The precondition for participation was a minimum age of 4 years, 0 months and a maximum age of 9 years, 0 months. Before running the study, we obtained parents' and/or a legal guardian's informed consent and ethics approval from the Ethics Committee of the Department of Economics and the Social Sciences at the University of Cologne. All methods were performed in accordance with the relevant guidelines and regulations of the ethics approval. Each session took about 30 minutes and was videotaped. The specific verbal instructions are available in the Appendix 4.C. The experiment took place around March and April 2018.

Children were randomly assigned to one of three between-subjects treatments: *Baseline*, *BetterWins*, and *WorseWins*. The test sessions were conducted with individual participants by a group of experimenters, trained after a detailed script. When investigating possible treatment effects, we pooled participants based on the median age (6.48) into two age groups: younger children (4.0-6.5 y/o) and older children (6.5-9.0 y/o), with no child aging 6.5 years. The respective sample sizes were: *Baseline*: 30 younger children, 27 older children; *BetterWins*: 27 younger children, 29 older children; *WorseWins*: 30 younger children, 29 older children. Two additional younger children were excluded from the analyses because they did not pass the control questions regarding understanding (see below). We use non-parametric statistical tests, e.g., Fisher's exact test and the Mann-Whitney U test. All reported tests are two-tailed.

4.2.2 Experimental Design

We employed a drawing task. In our main treatments, children chose one out of two (pre-selected) other drawings, which then served as their comparison standard in a relative evaluation of their own drawing. In particular, children proceeded through the following stages: (1) the drawing task, (2) questions of motivation and understanding, (3) the choice of other drawings (where instructions were partly treatment-specific), and finally (4) a self-control task.

Stage 1: Drawing task

In the drawing task, each child made three drawings from three sample photographs, depicting objects that were familiar to children and gender-neutral (a tree, a bird, a house, a giraffe, an owl, a bike, or the Cologne cathedral). Previous research has shown that drawings are typically of high relevance to children (Butler, 1989; Fu and Lee, 2007; Chevallier et al., 2012). For each kindergarten or school, children always received the same three randomly assigned photographs. Children had 3 minutes time to draw each of the photographs. The experimenter instructed children to draw as accurately as possible while a sandglass was running to keep time. At this point in time, children did not know about any treatment-specific rules.

Stage 2: Questions of motivation and understanding

The questions about motivation and understanding were asked once children had finished their drawings. Children were asked questions about (i) their desire to win in competitive games, (ii) their desire to be good at drawing, and (iii) their concerns about what others think of them (which we re-phrased for children as the desire to be 'liked by others'). Children had to indicate their level of motivation in response to these questions by pulling on a ruler (20cm length). Thereafter, children were asked questions of understanding concerning the notion of accuracy of drawings relative to a photograph. For this purpose, children received a folder containing two drawings of a sample photograph. Children were asked which of the two drawings was more accurate and which was less accurate in comparison to the photograph. Previous research has shown that children are able to assess the accuracy of drawings from about 3 years of age (Fu and Lee, 2007; Chevallier et al., 2012). In the treatments *BetterWins* and *WorseWins* (see below), children then had to point at the drawing that should win according to the rule explained by the experimenter. If children did not answer correctly for all of the questions of understanding, they entered another trial. Children went through a maximum of four trials using different materials, and they were excluded from the sample only if they failed in all four trials.[2]

Stage 3: Choice of other drawings

Next, children needed to make *three* subsequent choices. In each of their choices, they had to choose one of two other drawings. For each of their *three* own drawings from photographs,

[2]Note that children's understanding of the treatment-specific instructions, as indicated by the control questions (see Appendix 4.A), was very high also among the younger children: In each condition, at least 89% of the children responded correctly on the first trial and only 2 children had to be excluded because they did not manage to respond correctly within four trials.

children received a folder containing the two other drawings of the same photograph. The other drawings had been drawn by other children prior to the study. The experimenter explained that these drawings were made by other children and that the other children had the same amount of time and the same sample photograph to draw from. We pre-selected the other drawings such that the children's drawings were of middle accuracy. Importantly, children were only allowed to inspect the folder once the experimenter had his or her back turned, thus avoiding confounds due to the direct observation through the experimenter. The experimenter then explained to children that they would need to pick one of the two other drawings from the folder and lay it next to their own drawing. Before children actually picked a drawing, the experimenter explained the treatment-specific rules (see below, Subsection 4.2.3).

Stage 4: Self-control task

After children had chosen the other drawings, children proceeded to the self-control task. In particular, we administered the day-night task (Gerstadt et al., 1994), which is a stroop task customized to the targeted age group. Akin to the 'Marshmallow task' (see Mischel et al., 1989) or measures of economic time preferences (for a review see Sutter et al., 2019), it measures the ability to inhibit conflicting impulses. It has been shown to correlate positively with academic success and negatively with criminal behavior (Diamond, 2013). Children received a sheet with several pictures of suns and moons and needed to either inhibit the prepotent response of matching the word to the objects they see (incongruent trial, e.g., say "sun" whenever seeing a moon) or not (congruent trial, e.g., say "sun" whenever seeing a sun). For each type of trial, congruent and incongruent, we recorded how often children said the correct word within a set response time of 45 seconds (following the procedure used in, e.g., Archibald and Kerns (1999)). We used the number of correct utterances in the incongruent trial as a measure of children's level of inhibitory control.

4.2.3 Treatments

Before children chose one of the other drawings, they learned the rules of the evaluation procedure. In *Baseline*, the experimenter told children that s/he will look at the child's drawing and the other drawing after children had made a choice. Children received two gold tokens regardless of which drawing they had chosen from the folder. This treatment indicated children's basic preferences independent of any prize-winning incentives and thus allowed identifying which of the two other treatments would elicit an internal conflict (see also the

Hypotheses section). Further, in this treatment Stage 3 (Choice of other drawings) preceded Stage 2 (Questions of motivation and understanding), since children's understanding was less important in this treatment.

Contrary to *Baseline*, in *BetterWins* and *WorseWins*, whether or not children won a prize (i.e., received a positive evaluation) depended on the comparison standard that children selected. In *BetterWins*, children won, if the experimenter evaluated their drawing to be more accurate than the selected other drawing. In *WorseWins*, children won, if the experimenter evaluated their drawing to be less accurate than the selected other drawing. If children won the evaluation process in any of these two treatments, they received three gold tokens. If they did not win, they received 'only' one gold token. Note further that in both treatments, the experimenter always evaluated the children's drawing as more accurate only if the child chose the low-accuracy other drawing. At the end of the experiment, children could use the gold tokens they accumulated over the experiment to exchange for differing age-appropriate rewards (e.g., a marble, a roundabout, or a sports ball). In each of the treatments, we also asked children for their (qualitative) reasoning of why they had chosen a particular other drawing after children had made their last of three choices of the other drawings.

4.3 Hypotheses

To successfully engage in social-comparison engineering, children need to understand the potential consequences of social comparisons for their evaluations. For instance, children need to understand that one can achieve a better evaluation by choosing a lower comparison standard. While even three year-olds correctly understand and use social comparison statements, e.g., "My picture is better than yours." (e.g., Mosatche and Bragonier, 1981), children develop a basic understanding of the potential consequences of social comparisons for their evaluations only around late preschool age. Rather young preschool children often do not change their self-evaluations after experiencing relative failure or negative feedback (Butler, 1989; Ruble et al., 1976, 1980; Pomerantz et al., 1995; Ruble et al., 1994; Sutter and Glätzle-Rützler, 2015). Only from about 5 years of age, children start to rate their own performance less positively when another child has performed better than the child her/himself (Magid and Schulz, 2010; Butler, 1998; Rhodes and Brickman, 2008). Further studies have shown that older children of about 11 to 12 years of age tend to name classmates or friends as their comparison standards that enhance their evaluations (Tesser et al., 1984; Altermatt and

Pomerantz, 2005; Huguet et al., 2001). Somewhat related, research on children's strategic reasoning also suggests a development from about 7 years of age. This research investigates, for instance, whether children are able to think ahead a few moves (Brocas and Carrillo, 2018a,b), or whether they anticipate others' deceptive moves (Sher et al., 2014). We thus recruited children between the ages 4 and 9 to see whether also younger children would be able to select their comparison standards strategically in an age-appropriate situation.

Social-comparison engineering might be difficult for children (and even for adults) because, to enhance their relative evaluations, they have to overcome their preference for certain highly favored comparison standards. Hence, it was of high importance to pin down children's preferences over low-accuracy other drawings and high-accuracy other drawings to see where an internal conflict might occur. In this regard, Frey (1994)'s distinction between *basic* and *revealed* preferences is helpful, which he developed to describe people's preferences for art. Basic preferences refer to the "fundamental wishes of people and are taken as exogenously given" (p. 12), whereas revealed preferences are governed by external (economic or societal) constraints. In line with this definition, in the treatment *Baseline*, we aimed at eliciting children's *basic* preferences, that is, their preferences independent of any prize-winning incentives (i.e., economic constraints).[3] Treatment *Baseline* was thus exploratory as we did not have clear a priori hypotheses about what children's basic preference would be. We merely expected children to exhibit a preference for either low- or high-accuracy other drawings. Aesthetic considerations (Dutton, 2009), or similarity perceptions (Meer, 2011) might have played a role in this decision, but neither economic nor psychological theorists have addressed how aesthetic preferences come about in great detail. We interpret any preference different from pure random choice discovered in *Baseline* as an indicator for an internal conflict in one of the treatments with prize-winning incentives (*BetterWins* and *WorseWins*). Choosing against this preference might be like resisting a temptation (Gul and Pesendorfer, 2001; Fudenberg and Levine, 2006), which might be overcome if motivations to win are sufficiently strong or if self-control has developed to a certain extent (see below for hypotheses on the psychological processes). A development in children's ability for social-comparison engineering, i.e., age

[3]Of course, we cannot entirely exclude the presence of external constraints, since children make a choice which is subsequently observed: Children might have desired to "bask in reflected glory" or to liken themselves to "the better ones" by selecting a high-accuracy other drawings (Cialdini et al., 1976; Cialdini and De Nicholas, 1989; Collins, 1996). Alternatively, they might have desired to enhance their self-image as a good drawer or winner by selecting a low-accuracy other drawing (Tesser, 1988). Nevertheless, we contend that the elicited preferences are as close as we can get to basic preferences, even if they cannot be 'basic' in a strict sense. However, to measure whether or not children faced an internal conflict, it is irrelevant where precisely children's preferences originated in or whether they are 'basic' in a strict sense.

differences, should then occur in the treatment where children face an internal conflict, but not in the treatment where they do not face an internal conflict.[4] We sum up our expectations for the two different cases by differentiating in Hypothesis 1 (H) and Hypothesis 1 (L).

If children exhibit a basic preference for *high-accuracy other drawings*, children should face an internal conflict in the treatment *BetterWins*. Hence, we also expected age differences in social-comparison engineering, i.e, the ability to win the prize, to occur in this treatment:

Hypothesis 1 (H): *In BetterWins, younger children should win less often than older children (by choosing low-accuracy other drawings as their comparison standard). By contrast, children in the treatment WorseWins should win similarly often regardless of their age.*

If children exhibit a basic preference for *low-accuracy other drawings*, children should face an internal conflict in the treatment *WorseWins*. Hence, we also expected age differences in social-comparison engineering, i.e, the ability to win the prize, to occur in this treatment:

Hypothesis 1 (L): *In WorseWins, younger children should win less often than older children (by choosing high-accuracy other drawings as their comparison standard). By contrast, children in the treatment BetterWins should win similarly often regardless of their age.*

With respect to the psychological processes, we hypothesized that (i) children's desire for a positive evaluation outcome and (ii) their self-control should play a role in children's developing abilities for social-comparison engineering. A desire for a better evaluation outcome (i.e., winning) might arise either due to extrinsic (reward) or social image (reputation) concerns. Developmental research shows that a desire for positive evaluations is already present in young children. An observational study (Stipek et al., 1992) found that already 3-year-olds 'look away' after losing and express joy after winning in competitive and non-competitive tasks. Further, 4-5-year-old preschoolers (i) inflate their own performance relative to that of others (Ruble et al., 1994, 1976), and (ii) prefer tasks where they performed better (rather than worse) than others (Butler, 1998). Also the economic literature studying the development of competitive preferences in children (e.g. Sutter and Glätzle-Rützler, 2015) suggests that already children from 3 years expect to belong to the better ones (i.e., strong overconfidence), which is indicative of their desire to be better than other children and of their desire to earn rewards. Children from about 4 to 5 years of age are not only motivated to win, but they also

[4] Particularly among younger children, there might be differences in winning frequencies between the treatment where children face an internal conflict and the treatment where they do not face such a conflict. Note that however that in the treatment where children do not face an internal conflict, we cannot tell whether children are engaging in social-comparison engineering or whether they simply follow their basic preference (elicited in *Baseline*).

care about what others think, or equivalently, they want to be liked by others and have a good reputation. For example, young children share more with other children when peers or experimenters are watching (Engelmann et al., 2012; Shaw and Olson, 2014; Engelmann et al., 2013, 2016; Herrmann et al., 2019) than when nobody is watching. Further, they refrain from showing off, and they engage in flattery behavior (Banerjee, 2000; Banerjee et al., 2010; Fu and Lee, 2007; Bennett and Yeeles, 1990). The developmental literature thus suggests that even before children have a deeper understanding of social comparisons, they have a strong desire to win and care about what others think. Accordingly, even though children's desires for positive evaluations might predict their performance to some extent, these desires are unlikely to explain age differences in social-comparison engineering.

Importantly, we also expected social-comparison engineering to require some level of self-control because children might need to exert self-control over their own basic preferences (elicited in *Baseline*) to achieve positive evaluations. In the psychological literature, the two most important components of self-control (or self-regulatory abilities) relate to children's working memory, i.e., the ability to keep goal-relevant information in mind, and inhibitory control, i.e., the ability to resist temptations and to not act impulsively (Diamond, 2013). Inhibitory control has been shown greatly benefit people's school and health outcomes (Moffitt et al., 2011; Diamond, 2013). Inhibitory control increases around the age of 4-6 years (Blair and Razza, 2007), and it improves further in later years (Zelazo and Carlson, 2012; Archibald and Kerns, 1999). Particularly inhibitory control might contribute to explaining children's developing abilities for social-comparison engineering, because they need to resist the temptation to choose the preferred other drawings. Hence, inhibitory control might explain age-related differences in children's winning frequencies in the treatment where they face an internal conflict, i.e., their ability to select the prize-winning comparison standard. Hypothesis 2 sums up our expectations regarding the psychological processes.[5,6]

Hypothesis 2: *In the treatment where children face an internal conflict, their increasing level of inhibitory control, but not their desire to win or their desire to be liked by others, should explain their ability to win by selecting the prize-winning comparison standard.*

[5]Inhibitory control measures facets that are related to time preferences. Mischel et al. (1989)'s seminal study on the delay of gratification showed that the performance of 4 year old children in the marshmallow task predicted their social and cognitive competence as adolescents. Further, economic time preferences, like inhibitory control, increase between the ages 3 and 6 (Sutter et al., 2015), and are also important to predict field behavior, e.g., drug misuse (Sutter et al., 2013).

[6]Note that we also elicited children's desire to be good at drawing as a measure of their motivation to participate in our task, but we did not expect children's drawing motivation to vary with age based on previous work, e.g. by Butler (1989), or explain any of the age differences observed.

4.4 Results

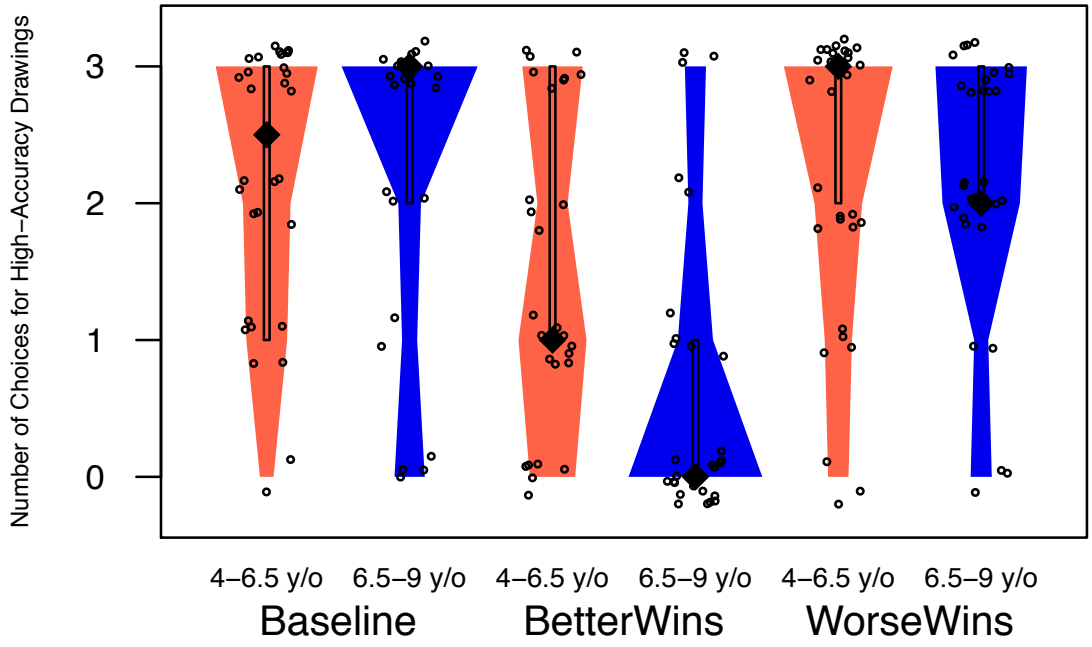


Figure 4.1: Number of trials in which children chose the high-accuracy other drawing (out of three) as a function of treatment and age group. The two age groups 4-6.5 y/o and 6.5-9 y/o were constructed by splitting the sample at the age of 6.5 (median 6.48). Circles show individual observations and violins show rotated histograms. Diamonds represent medians and boxes represent the interquartile ranges, i.e., the range between the 25$^{\text{th}}$ and the 75$^{\text{th}}$ percentile.

Figure 4.1 shows the distributions and medians of the number of choices (out of 3) for the high- vs. low-accuracy other drawings as a function of treatment and age group.[7] The figure shows that, in the *Baseline* treatment where the selection of the other drawing had no material consequences, children exhibited a clear preference for the high-accuracy other drawing (4-6.5 years old: mean 2.17 (out of 3 trials, i.e., 72% of the trials), SD = 0.99; Fisher's test against chance, p = .016; 6.5-9 years old: mean 2.30 (77%), SD = 1.14; Fisher's test against chance: $p < .001$). This preference was regardless of children's age (Mann-Whitney U test as for the difference between age groups: Z = -0.87, p = .399, r = 0.11, Spearman's rank correlation between number of high-accuracy drawings and age, ρ = .018, p =0.893). Result 1 sums up the findings regarding children's basic preferences.

[7]Note that for illustration purposes, two age groups 4-6.5 years old and 6.5-9 years old were constructed by splitting the sample at the age of 6.5 (median 6.48), with no child aging exactly 6.5 years. In our statistical analyses below, we report non-parametric tests between different age groups *and* complementing rank correlation tests using age as a continuous variable.

Result 1: *Regardless of their age, children exhibited a strong preference for high-accuracy other drawings in the Baseline treatment.*

Given the clear preference for the high-accuracy other drawing in *Baseline*, in line with Hypothesis 1 (H), children should experience no internal conflict in *WorseWins*, where their own drawing had to be less accurate to win the prize, i.e., choosing high-accuracy other drawings made them win the prize. In contrast, in *BetterWins*, a preference for the high-accuracy other drawing conflicted with choosing the prize-winning action, i.e., a low-accuracy other drawing. In line with this assumption, in *WorseWins* children predominantly chose the high-accuracy other drawing, regardless of their age (4-6.5 y/o: mean 2.20 (73%), SD = 1.03; Fisher's test against chance: p = .009; 6.5-9 y/o: mean 2.17 (72%), SD = 0.97; Fisher's test against chance: p = .008; Mann-Whitney U test as for the difference between age groups: Z = -0.31, p = .771, r = 0.04, Spearman's rank correlation between winning frequencies and age, ρ = .103, p = 0.445). By contrast, in line with Hypothesis 1 (H), in *BetterWins*, when an internal conflict was likely, children's performance increased with age and their behavior substantially differed from *Baseline* and *WorseWins*. Notably, younger children faced difficulties in choosing the prize-winning comparison standard: They chose both the high-accuracy and the low-accuracy other drawing similarly often in *BetterWins* (mean 1.52 (51%), SD = 1.16, Fisher's test against chance: p = .180). In contrast, older children only rarely chose the high-accuracy other drawing in this condition (mean 0.62 (21%), SD = 1.02; Fisher's test against chance: p = .009).[8] In direct support of Hypothesis 1 (H), younger children won less often than older children in *BetterWins* (Mann-Whitney U test as for the difference between age groups: Z = 3.12, p = .001, r = 0.42, Spearman's rank correlation, ρ = 0.344, p =.009), and there was no similar age increase in the treatment *WorseWins* (Mann-Whitney U test as for the difference between age groups: Z = -0.31, p = .771, r = 0.04, Spearman's rank correlation between winning frequencies and age, ρ = 0.018, p =.893). Also in support of Hypothesis 1 (H), younger children won less often in *BetterWins* (mean 1.48 (49%), SD = 1.16) than in *WorseWins* (mean 2.20 (73%), SD = 0.97; Mann-Whitney U test as for the difference in winning frequencies: Z = 2.45, p = .015, r = 0.33), whereas older children won equally often in *BetterWins* (mean 2.38 (79%), SD = 1.02) and *WorseWins* (mean 2.17 (72%), SD = 0.97; Z = -1.24, p = .232, r = 0.16). Thus, younger children faced greater difficulties to achieve positive evaluations in social-comparison situations,

[8]Nevertheless, both younger and older children showed a choice behavior significantly different from that in *Baseline*, and chose the low-accuracy other drawing that maximized prizes significantly more often compared to *Baseline* (4-6.5 y/o: Z = 2.15, p = .033, r = 0.29; 6.5-9 y/o: Z = 4.55, $p < .001$, r = 0.61). This indicates that even though younger children faced great difficulties in choosing the prize-winning comparison standard, their behavior was affected by the prize-winning incentives to some extent.

particularly when there was an internal conflict. Children's qualitative reasoning, which we assessed after their choices, further supported our interpretation of children's choice behavior. Children in *Baseline* referred to the beauty of the high-accuracy other drawing (without making a comparison statement) as the dominant reason for their choice (in 39% of the cases). Children in the treatments with prize-winning incentives, in contrast, referred more to the quality of the other drawing relative to their own drawing (in 57% of the cases in *WorseWins* and in 50% of the cases in *BetterWins*; for detailed results, see Appendix 4.B). Result 2 sums up the findings regarding Hypothesis 1 (H).

Result 2: *When children faced an internal conflict in BetterWins, older children won significantly more often than younger children (by choosing a low-accuracy other drawing). By contrast, children won similarly often regardless of their age group, if they did not face an internal conflict (in WorseWins).*

Next, we investigated the psychological processes which might explain the age differences found in children's social-comparison engineering in *BetterWins*. We did not find any meaningful differences between younger and older children in the motivation to win (12.2 vs. 10.5, respectively), and the motivation to be liked by other children (14.0 vs. 13.6, respectively) (Mann-Whitney U tests as for the difference between age groups, all p-values $> .10$, Spearman's rank correlation between the motivations and age, $\rho \approx 0$, p-values $> .10$). Yet, we did find an age-related difference in children's inhibitory control as indicated by their performance in the age-appropriate day-night stroop-task (4-6.5 y/o: mean 26, $SD = 9$; 6.5-9 y/o: mean 39, $SD = 8$; Mann-Whitney U test as for the difference between age groups: $Z = 4.71$, $p < .001$, $r = 0.63$, Spearman's rank correlation between inhibitory control and age, $\rho = .620$, p-values $< .001$). Based on these observations, we investigated whether older children's increased level of inhibitory control, their desire to win, or their desire to be liked by others explained their higher likelihood to win, when there was an internal conflict in *BetterWins*. To this end, in Table 4.1, we report several multi-level linear probability models predicting children's likelihood to win in a given trial (out of three). All our models account for the dependencies at the level of children's daycare facilities and across children's trials.

In Model 1, we only include children's age group as a dummy variable indicating whether or not children were 6.5 to 9 years old. In this model, we again find support for the age difference documented in previous non-parametric tests ($b_{\text{6.5-9 y/o}} = 0.299$, SE $= .097$, $p = .002$). In Model 2, we do not account for children's age group, but only include their motivation to win, their

	Win		
	(1)	(2)	(3)
6.5-9 y/o	0.299***		0.281**
	(0.097)		(0.121)
Motivation to win		0.011	0.009
		(0.008)	(0.011)
Motivation to be liked by others		−0.007	0.006
		(0.008)	(0.010)
Inhibitory control		0.009*	−0.014*
		(0.005)	(0.008)
6.5-9 y/o* Motivation to win			0.006
			(0.014)
6.5-9 y/o*Motivation to be liked by others			−0.020
			(0.014)
6.5-9 y/o*Inhibitory control			0.029**
			(0.012)
Constant	0.494***	0.645***	0.421***
	(0.070)	(0.052)	(0.080)
Observations	168	168	168
Log Likelihood	−95.535	−108.522	−113.354
Note:			$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 4.1: Linear probability models predicting children's likelihood to win in treatment *BetterWins*.

motivation to be liked by others, and their inhibitory control. This model sheds light on the extent to which the motivational and cognitive variables explained children's likelihood to win. In Model 2, inhibitory control, but not children's motivation to win or their motivation to be liked by others, systematically predicted their likelihood to win ($b_{\text{Motivation to win}} = 0.011$, SE = .008, $p = .143$, $b_{\text{Motivation to be liked by others}} = -.007$, SE = 0.008, $p = .339$, $b_{\text{Inhibitory control}}$ = .009, SE = 0.005, $p = .057$). Model 3 extended Model 2 by also accounting for children's age. In particular, we interacted children's age group (whether or not they were 6.5-9 years old) with their desire to win, their desire to be liked by others, and their inhibitory control, respectively, and thus estimated how the relationships between our motivational and cognitive

variables differ by age. The younger children (4-6.5 y/o) were the reference category in this model, so the coefficient on *Inhibitory control* measured the extent to which higher inhibitory control increased the likelihood to win in younger children (analogously for *Motivation to win* and *Motivation to be liked by others*). The interaction between age group and *Inhibitory control*, by contrast, measured the extent to which higher inhibitory control increased the likelihood to win in older children (6.5-9 y/o) (analogously for *Motivation to win* and *Motivation to be liked by others*). Interestingly, supporting Hypothesis 2, we find that while the motivation to win and the motivation to be liked by others did not explain children's likelihood to win in either younger or older children (all coefficients' p-values $> .10$), higher inhibitory control significantly increased the likelihood to win in older children ($b_{\text{Inhibitory control*6.5-9 y/o}} = .029$, SE $= .012$, $p = .015$).[9]

Result 3: *When there was an internal conflict in BetterWins, older children's, but not younger children's, increasing level of inhibitory control explained their ability to choose the prize-winning comparison standard.*

4.5 Discussion and Conclusion

Our findings suggest that the ability to strategically select social comparison standards in order to achieve positive evaluations, even against the own basic preferences, develops in middle childhood. In our experiment, children of all ages showed a strong preference for choosing other drawings of high accuracy when there was no evaluation (*Baseline*). Likewise, if the choice of other drawings was followed by an evaluation in which the child won a prize if their own drawing was less accurate than the chosen drawing (*WorseWins*), children of all ages managed to win the prize similarly often. From just looking at this comparison, however, it remains unclear whether children simply followed their preferences or indeed strategically chose the high-accuracy other drawing in this condition. The more telling treatment is the one in which children faced the evaluation in which they would win a prize for their drawing being the more accurate drawing (*BetterWins*). Here, children had to inhibit their basic preference for other drawings of high accuracy in order to receive a positive evaluation. Only older children (6.5-9 y/o) were able to engineer the social-comparison situation when facing an internal conflict, outperforming younger children (4-6.5 y/o).

We also investigated the mechanism underlying children's social-comparison engineering:

[9]The reported statistical patterns remain unchanged, if we use children's age (in years) instead of their age group as a variable.

While the responses in children's motivation to win or to be liked by others did not substantially vary by age, children's inhibitory control increased significantly between 4 and 9 years of age. Further, increased levels of inhibitory control in the older children, but not in the younger children, explained children's increasing ability to select the prize-winning comparison standard. This finding suggests that inhibitory control needs to have developed to a sufficient degree such that it can support successful social-comparison engineering.

These findings have broad implications for our understanding of the developmental processes underlying social-comparison engineering as well as children's ability to strategically affect others' payoff-relevant evaluations. Previous research has shown that self-control is important for mental and physical health or academic success (Moffitt et al., 2011; Blair and Razza, 2007; Diamond, 2013) as well as for individual decision-making, e.g., in the marshmallow task or gambling tasks (Mischel et al., 1989; Garon, 2016). Our results broaden this perspective and suggest that self-control is also important for (the development of) acting strategically in social situations, and, more generally, when facing payoff-relevant evaluations.

To conclude, we provide clear evidence that social-comparison engineering undergoes a significant development. The ability to inhibit conflicting impulses provides children with the opportunity to strategically shape social comparisons made by others, and thus, to increase their success and, potentially, also their well-being.

Appendices

4.A Children's answers to questions of understanding

Age group	[4.0y-6.5y)			[6.5y, 9y]		
Treatment	Base	BW	WW	Base	BW	WW
N	30	27	30	27	29	29
Male	53%	56%	53%	44%	48%	55%
Q: 1st Trial	100%	89%	80%	96%	97%	76%
Q: 2nd Trial	-	89%	97%	100%	100%	100%
Q: 3rd Trial	-	96%	100%	-	-	-
Q: 4th Trial	-	100%	-	-	-	-

Table 4.2: Percentage of children who correctly answered all the control questions of understanding (Questions 5-7 in the Verbal Instructions). Two additional kids aged 4.5 and 6.2 failed at all trials and are excluded.

4.B Reasons for children's choices

Treatment	Baseline	WorseWins	BetterWins
Choice of other drawing	High-accuracy	High-accuracy	Low-accuracy
n	23	14	14
Reasons			
„it's beautiful“	39.1%	21.4%	14.3%
„it looks more like the photo“	30.4%	57.1%	
„it looks like mine“	21.7%		
„mine is better“			50.0%
„this yields more tokens“		21.4%	21.4%

Table 4.3: Children's responses when asked for their reasons of why they chose a drawing after they had chosen their last other drawing. Answers were categorized by pre-specified categories. Reasons for choices of other drawings were too low to classify for the forllowing cases: Baseline, low-accuracy drawings (n=9); WorseWins, low-accuracy drawings (n=7); and BetterWins, high-accuracy drawings (n=7).

4.C Instructions

Drawings from photographs

Greeting

Hello [name of the child]! I am [name of confederate, typically a female student of psychology or education who received a training to complete the interview].
We will soon play games where you can earn tokens.

First, we will draw, and then we will play the games with the tokens.
You can later exchange your tokens for one of the following prizes. [show prizes]
Try your best, so you earn lots of tokens!

Draw
Now we will draw. On the whole, we will make three drawings. You will later need the drawings in the games, where you can win the tokens.

Look, here is a pencil, a binder, a paper and a photograph. When I tell you, you draw the photograph. Take care that your drawing looks like the photograph. We know it's not easy, but try your best.

You have three minutes. You start when I turn the sandglass and the time is over when all sand has run down the sandglass.
Some rules: Tracing is forbidden. You may not only use this pencil and no rubber. Do you have any questions?

You can now start to draw the photograph.
It's not about being done first. Take care that your drawing looks like the photograph.

→ *[turn the sandglass]*
[after 1 minute] Now one minute has passed. You still have two minutes.
[after 2 minutes] Now two minutes have passed. You still have one minute. Now your drawing should be finished soon.
[nach 3 Minuten] Please put your drawing into your binder.
[no comments or praise at this point!]

[hand out new paper and new photograph]

Now you receive a new paper and a new photograph to draw from.
As I told you before, you will need the drawings for the games where you can earn the tokens.
You can now start drawing.

Again, the time will be over when all sand has run down the sandglass.
And remember to take care that your drawing looks like the photograph.
[3 drawings in total, continue at →]

We are done with drawing. *[continue with the games]*

Treatment BetterWins

Part 1: Questions of understanding

We will soon play the games where you can earn the tokens. Before we start, there are some questions. Look, here you have a ruler [printed on paper]. You can use it to show me how much you want something.

If you make the ruler very short, you do not want something very much.
If you make the ruler very long you, you want something very much.
[show very short 1cm, show very long 19cm]

The longer you make the ruler, the more you want something.

1. How much do you want to be good at drawing?
2. How much do you want to win in games, where other kids are also participating?
3. How much do you want that others like you a lot?

Now we are playing the first game. I will show you a folder. This folder has two drawings in it which other kids drew who are not from your group.
[open the folder containing the drawings]

→ The other kids tried to draw this photograph.

4. What do you think, which of the two drawings should win a prize?
5. Which of the two drawings looks more like the photograph?
6. Which of the two drawings looks less like the photograph?
7. Let's make a rule:

   The drawing wins the prize that looks MORE like the photograph. Which drawing has to win the prize?

[If question 5-7 were correct, continue with Part 2]

[If not, please correct (as below) and pull out new folder with new photograph and ask questions 5-7; repeat until child has seen a maximum of four folders]

This drawing looks more like the photograph. This drawing wins the prize.

Now we will do the same with this folder. Again, this folder has two drawings in it that were drawn by other kids who are not from your group. [continue above at →]

Part 2: Choosing between drawings

Now we will play the game where you can win the tokens. You will receive one folder for each photograph that you drew.
[show, but do NOT open folder!]

The folder has two drawings in it which other kids drew who are not in your group. The other kids tried – just like you – to draw such that their drawing looks like the photograph. The other kids also had three minutes.

Before we start: You may only look at the drawings in the folder when I have turned around. After this you have to close folder again.

Now it's about this folder. This means it's about THIS photograph [point]. This is YOUR drawing of the photograph.

→ When I turn around, then you will lay ONE drawing from the folder NEXT TO your drawing.

When you are done, I will decide, whether your drawing looks more like the photograph or whether the other drawing which you laid next to your drawing looks more like the photograph.

If I decide that your drawing looks more like the photograph, you win THREE tokens.

If I decide that your drawing looks less like the photograph, you win ONLY ONE token.

Alright? Afterwards, the two drawings will be put on this board here.

So, I am now turning around. Please lay ONE drawing of the other kids next to your drawing. Have you closed the folder? *[turn back again]*

[Only if the child chooses the drawing that looks less like the photograph, the drawing of the child is evaluated as looking "more like the photograph"]

[Positive:] Well, your drawing looks more like the photograph. Here you have three tokens.

[Negative:] Well, the other drawing looks more like the photograph. Here you have one token.

[the chosen drawing will be attached to the board together with the child's drawing]

Now we will do the same with the other folders. This is the second folder. This means it's about THIS photograph. This is YOUR drawing of the photograph. [continue above at →]

[repeat until all three folders complete]

Why did you choose this drawing?
Very well done, you have already earned some tokens! I think your drawings are great.

Do you have questions regarding your drawings or the drawings from the folders?
Do you have questions about what I said regarding your drawings or the drawings from the folders?

[Who? An older child that is good at drawing; explain evaluations as part of the game and the rules]

Now there is a different game where you can win tokens.

Treatment WorseWins

For WorseWins, two paragraphs were different from BetterWins:

Part 1, Question 7 "The drawing wins the prize that looks LESS like the photograph"

Part 2: Choosing between drawings (grey area) is now:

When I turn around, then you will lay ONE drawing from the folder NEXT TO your drawing.

When you are done, I will decide, whether your drawing looks more like the photograph or whether the other drawing which you laid next to your drawing looks more like the photograph.

If I decide that your drawing looks more like the photograph, you win ONE token.

If I decide that your drawing looks less like the photograph, you win THREE tokens.

Alright? Afterwards, the two drawings will be put on this board here.

So, I am now turning around. Please lay ONE drawing of the other kids next to your drawing. Have you closed the folder? *[turn back again]*

[Only if the child chooses the drawing that looks less like the photograph, the drawing of the child is evaluated as looking "more like the photograph"]

[Positive:] Well, the other drawing looks more like the photograph, here you have three tokens.

[Negative:] Well, your drawing looks more like the photograph, here you have one token.

Treatment Baseline

Different from BetterWins, for Baseline, Part 1, Question 7 was not asked. Further, Part 1: Questions of understanding was asked only after Part 2: Choosing between drawings had been completed not to confound children's baseline preference.

Part 2: Choosing between drawings (grey area) is now:

When I turn around, then you will lay ONE drawing from the folder NEXT TO your drawing.

When you are done, I will look closely at your drawings and the other drawing which you laid next to your drawing.

Alright? Afterwards, the two drawings will be put on this board here.

So, I am now turning around. Please lay ONE drawing of the other kids next to your drawing. Have you closed the folder? *[turn back again]*

[Only if the child chooses the drawing that looks less like the photograph, the drawing of the child is evaluated as looking "more like the photograph"]

[*Feedback*:] Here you have two tokens.

Part 3: Sun-moon task
Now we will play the last game.
[pull out sheet with suns and moons and control sheet]

Look here you have a sheet. Look, there is this picture. If you see this picture, I want that you say "sun".
[point at sun in the upper left corner]

Can you repeat the right word for the picture?

Here you see a different picture. If you see this picture, I want that you say the word "moon".
[point at moon in the upper left corner]
Can you repeat the right word for the picture?

[repeat this procedure until the child has identified both pictures correctly.
If the child did not succeed after four trials, do not play the game.]

Your task is to say as often as possible the correct words for the pictures. You begin in the upper row. In each row you then move from left to right. Then you continue one row below. If you are done with the sheet, you once again start in the upper left.
[show sequence by pointing]

When I say „GO", you have 45 seconds. Alright? "GO"

Look here you have a last sheet. Look, there is this picture. If you see this picture, I want that you say "moon".
[point at sun in the upper left corner]

Can you repeat the right word for the picture?

Here you see a different picture. If you see this picture, I want that you say the word "sun".
[point at moon in the upper left corner]
Can you repeat the right word for the picture?

[repeat this procedure until the child has identified both pictures correctly.
If the child did not succeed after four trials, do not play the game.]

Your task is to say as often as possible the correct words for the pictures. The sequence is as before.
[show sequence by pointing]

When I say „GO", you have 45 seconds. Alright? "GO"

We are done. It was great how you participated. Thank you so much! You earned 6 tokens in this last game. Now you can choose your prize.

You will receive the prize and your folder with the drawings only when all children have finished.

Chapter 5

The desire to set an example

Do people cooperate more if they have social influence?

5.1 Introduction

A recent meta-study (Bradley et al., 2018) demonstrated that people are more likely to cooperate, when they feel observed by others, so potentially their social image is at stake. However, when people feel observed by others, they might also realize that their decisions will provide social information for others. Social information is known to influence diverse social decisions, such as donations or public goods contributions (Chen et al., 2010; Frey and Meier, 2004; Shang and Croson, 2009). Hence, people might anticipate the social influence of their decision on others' decisions and thus perceive an opportunity to 'set an example' for others. For instance, if an individual engages in a local social project, he or she might be motivated by the idea of inspiring others to also more strongly engage in their social projects. Previous studies on cooperation have investigated situations where people could set an example for others, e.g., for another group (Chaudhuri and Graziano, 2006), or for their own group ('leading-by-example', e.g., Güth et al., 2007). In these studies, however, people faced monetary incentives to set an example. Hence, it is still an open question whether or not humans more generally desire to set an example for others, i.e., also in the absence of monetary incentives, which is interesting for our understanding of social learning and the evolution of cooperation (e.g., Boyd et al., 2011; Tabellini, 2008). In addition, for charitable organizations or local governments aiming to encourage prosocial behavior, it is important to understand *why* people set an example for others. Possibly, people's desire to set an example arises merely due to their prosocial motivations, e.g., due to their preference for fostering efficient outcomes (e.g., Charness and Rabin, 2002). In this case, to appeal to people's

desire to set an example, it might suffice to make their contributions public, so their decisions can influence others' social information. Alternatively, based on the 'focusing theory' of social norms (Cialdini et al., 1991), people's desire to set an example might arise only when the normative aspects of a decision situation become sufficiently salient, so people's self-image as a prosocial person is at stake (Bem, 1972; Leary and Baumeister, 2000; Bénabou and Tirole, 2011). From this perspective, to appeal to the desire to set an example, people's attention should be focused on injunctive norms, e.g., by giving them an option to personally recommend to others what *should* be done. In previous studies (Gächter et al., 2012; Arbak and Villeval, 2013), inferences about the psychological motives for setting an example were complicated by monetary incentives, and they focused on correlations between measures of prosocial motivation and contributions. In this paper, I therefore investigate experimentally whether or not - and why - people set an example for others' cooperation decisions in the absence of monetary incentives for doing so.

In the experiment, participants had social influence on others' cooperation decisions, so they could set an example for others without monetary incentives. Participants were either assigned to 'first-teams' or to 'follower-teams' to play a two-player public goods game. Each player in a first-team provided information about his/her contributions for the players of his/her own 'follower-team'. To exclude monetary incentives for setting an example for follower-teams, each team played its own public goods game independently of the other teams. To vary whether or not a first-team player had social influence on his or her follower-team's contribution decisions, follower-team players either had the social information available before they made their contribution decisions (*Infl* treatments), or only *after* they had made binding contribution decisions (*NoInfl* treatments). The *NoInfl* treatments thus controlled for social image motives due to mere observability of the first-team player's contribution (Bradley et al., 2018). To vary the salience of injunctive norms (and thus self-image motives to appear prosocial), first-team players either had an opportunity to write a recommendation about what their respective follower-team players *should* contribute (*Rec* treatments), or they did not have such an opportunity (*NoRec* treatments).

The findings in this study highlight that, if first-team players had social influence and could write a recommendation, first-team players contributed significantly more to their team's public good than first-team players who did not have social influence and could (or could not) write a recommendation. Additionally, they also reported a higher motivation to set an example. By contrast, for first-team players who had social influence but could not write a recommendation, there was no increase in contributions relative to the *NoInfl* control treatments. People's desire

to set an example thus depended on the salience of injunctive norms, which, in turn, activated their desire for positive self-image as prosocial persons.

As a second contribution, this paper sheds light on the impact of social information on 'conditional cooperation', i.e., the tendency to contribute if others do so (Kelley and Stahelski, 1970; Fischbacher et al., 2001), which is important to understand how policy makers can employ social information to increase cooperation. In particular, the design allows investigating how the social information provided by first-team players affected contributions of the follower-teams that received the information. One might expect that social information about high (or low) contributions should simply increase (or decrease) follower-team player's beliefs about their co-player's contributions and thus also their contributions. Yet, I document that follower-team players were more likely to follow high first-team player contributions than they were to follow low first-team player contributions. This suggests that, in addition to affecting people's beliefs, the social information might have focused people on the normative aspects of the decision situation (Cialdini et al., 1991; Krupka and Weber, 2009), making it more attractive to follow high (normative) contributions.

5.2 Previous Literature

Presious research in the public goods literature focused on leading-by-example in a sequential public goods game (Güth et al., 2007; Levati et al., 2007; Arbak and Villeval, 2013; Eckel and Wilson, 2007; Gächter et al., 2012; Potters et al., 2007; Rivas and Sutter, 2008, 2011; Gächter and Renner, 2018; Gürerk et al., 2018). In the leading-by-example setting, there are monetary incentives to set an example: Cooperation is the money-maximizing behavior, if a leader believes that his or her group members are conditional cooperators. Research investigating social learning in public goods games or other games focused on the effect of advice on the cooperation decisions of subsequent 'laboratory generations' (Chaudhuri and Graziano, 2006; Schotter and Sopher, 2003, 2006). Also in these studies, there were monetary incentives for giving advice, since a large share of people's monetary payoffs depended directly on the payoffs of the subsequent 'laboratory generations' that received their advice. In fact, most of the literature on advice focuses on the effects of incentivized advice (e.g., Kuang et al., 2007; Danilov et al., 2013). In a notable exception, Schram and Charness (2015) investigated the effect of moral advice on dictator game giving, where the particular advice given could not decrease or increase people's monetary payoffs. Yet, the players who gave advice in the Schram and Charness (2015) study

did not make any decisions themselves, so, contrary to the present study, they did not have social influence on others via their own economic decisions.

There is also a literature on the impact of social information in donation settings (Croson and Shang, 2008; Frey and Meier, 2004; Krupka and Croson, 2016; Shang and Croson, 2009). In these studies, individuals simply received social information which originated in the donations of previous individuals who were unaware that their decisions would later provide information for others. Perhaps most closely related to the present study, Reinstein and Riener (2012) presented a donation game, where players were aware that their donations would provide information for others. Before the experiment started, participants went through a meet-and-greet-stage, so people's social image outside the lab became salient. The donations were highest in a treatment where people had their identity and information about their donations revealed. Yet, the authors did not find any significant differences with respect to their relevant control conditions, possibly because people's social image outside the lab became too salient, because social influence itself was weak in their setting, or simply because the study had a small sample size of $N = 24$ per treatment. Previous studies thus did not isolate people's desire to set an example for others independent of monetary incentives for doing so, and they did not differentiate experimentally between different motives for doing so – which is the goal of the present study.

5.3 Experiment

5.3.1 Participants and Procedure

The data was collected via the online platform Amazon Mechanical Turk during weekdays. Note that a recent meta-study (Bradley et al., 2018) on the effect of mere observability on prosocial behavior has revealed rather weak observability effects, especially in online settings. Using an online setting is thus helpful to avoid that strong social image motives due to mere observability divert people's attention from the opportunity to set an example. To ensure that participants read and understood the instructions, participants had to answer three control questions correctly in order to be allowed to participate and receive payments. Participants could further make use of a payoff calculator with sliders. For each question, subjects needed to mark the correct statements from a set of four or five statements. The first question addressed subjects' understanding of the public goods game, the second addressed their understanding of the game setup and the third addressed their understanding of the belief elicitation procedure. The sample consists of 1440 participants, resulting in 120 independent observations, i.e., 120 first-team players and 240

follower-team players, for each of the four treatments.[1]

5.3.2 Experimental Design

Participants were randomly matched into different teams to play the public goods game. Within each team, participants played a simultaneous two-player public goods game. Each player received an endowment of \$1.00, of which he or she could contribute a multiple of 10 Cents to his or her team's common project. Each Cent contributed was multiplied by 1.5 and distributed equally among the two players. The marginal-per-capita-return (MPCR) to cooperation was thus set rather high (0.75) which should ensure that cooperative behavior was (perceived as) the norm. Empirically, a high MPCR facilitates cooperation, even though free-riding incentives remain unchanged (Isaac and Walker, 1988).

To create a situation where some players can have social influence on others, players were either assigned to the 'first-team' (labeled Pair A in the instructions) or to one of two 'follower-teams' (labeled Pair B or Pair C in the instructions). The setup thus consisted of one first-team and two follower-teams. Players in the first-team played first and, subsequently, the players of the two follower-teams played. To exclude monetary incentives to set an example, each team played its own game separately from the other teams and thus could not directly affect the monetary payoffs in other teams. One player in the first-team provided information for the two players of one follower-team, and the other player in the first-team provided information for the two players of the other follower-team. Each first-team player thus had 'exclusive' social influence over two follower-team players. This should ensure that first-team players perceived their contribution to be sufficiently influential and that conflicting information from a co-player could not weaken a first-team player's social influence. To further ensure that first-team players' contributions indeed affected follower-team players' contributions, players of the same follower-team were informed that they would both receive the same information from the same player in the first-team, and first-team players knew this. Hence, the social information that the first-team players provided became public knowledge, giving the first-team players an opportunity to boost conditional cooperation in follower-teams (Fischbacher et al., 2001). To further exclude any form of 'social competition', i.e., the desire to be more prosocial than others (e.g. Böhm et al., 2013), the first-team players were only informed about the contributions of their own follower-team players and not about the contributions of the other follower-team. Alike, follower-team players

[1] Of the 2287 recruited participants, 389 dropped out and 458 were excluded by the program because they did not answer correctly one of the control questions within three trials. Importantly, the number of participants who dropped out or were excluded did not vary significantly by treatment (Chi-Square test, $X^2 = 2.80$, $p = 0.423$).

of the same team did not receive information about the contributions of the other follower-team.

5.3.3 Treatments

The 2x2 treatments varied (i) whether or not the first-team players had social influence (*Infl* treatments) or not (*NoInfl* treatments), and (ii) whether first-team players could (voluntarily) write a recommendation about what should be done (*Rec* treatments) or not (*NoRec* treatments).

The *Infl/NoInfl* variation allows isolating the effect of having social influence on the desire to set an example separately from social image motives or any other confounding motives that might occur if people know their decisions will be made available to others. In the treatments *Infl-Rec* and *Infl-NoRec*, the follower-team players receive information about their first-team player's contribution *before* they make their contributions to their team's project. Hence, first-team players have social influence on their follower-team players' decisions via their contribution decisions. By contrast, in the treatments *NoInfl-Rec* and *NoInfl-NoRec*, the follower-team players receive the information about their first-team player's contribution only after they have already finished playing their team's public goods game, i.e., after they have made binding contribution decisions. These treatments thus control for social image motives, e.g., due to the feeling of being observed by the follower-team players.

The *Rec/NoRec* variation allows isolating the impact of the salience of injunctive norms on people's desire to set an example. In the treatments *Infl-NoRec* and *NoInfl-NoRec*, first-team players merely provide information about their contribution and an opportunity to write a recommendation for their follower-team is not mentioned. As a result, injunctive norms are not particularly salient. By contrast, in *Infl-Rec*, each first-team player can voluntarily write "a recommendation about what the players in [his or her follower-team] should contribute". This manipulation thus makes first-team players think about what should be done, e.g., the injunctive norm to make a high contribution. In the *NoInfl-Rec* treatment, first-team players also had an opportunity to voluntarily write a recommendation. Yet, first-team players knew that their contribution and their recommendation (if provided) was displayed to follower-team players only after they had already finished making their decisions.[2] Hence, like darkness makes salient the lack of observability (Zhong et al., 2010), this treatment might make salient the irrelevance of first-team player's decision for follower-team player's decisions.

To control for the effect of the treatments on beliefs, I also elicited players' beliefs. The

[2]In the treatment *NoInfl-Rec*, first-team players' recommendation is labelled "comment about what the players in [his or her follower-team] should have contributed" to make clear that it cannot technically affect their follower team players when they make their contribution decisions.

belief elicitation took place only after the players had made their contribution decisions to avoid confounding players' contribution decisions (Blanco et al., 2014). In all treatments, first-team players and follower-team players guessed the average contributions of all first-team players. Further, first-team players in the treatments *Infl-NoRec* and *Infl-Rec* had social influence on their follower-teams' decisions, so it was important to find out how participants expected first-team player contributions to affect follower-team player contributions. To this end, first-team players and follower-team players each guessed the average contributions of players in follower-teams for the cases of receiving information about a 0, 50, and 100 Cents contribution by first-team players. First-team players in the treatments *NoInfl-NoRec* and *NoInfl-Rec* did not have social influence on their follower-teams' decisions. Accordingly, in these treatments, first-team players and follower-team players each guessed the average contribution of all follower-team players without conditioning on information about different contribution levels.

In the belief elicitation, players could only submit multiples of 10 Cents as guesses and these guesses were compared to the actual average contribution rounded to the nearest multiple of 10 Cents of the respective data collection days. Players' guesses were incentivized by a simplified quadratic scoring rule: If their guess was exactly equal to the (rounded) average contribution, their payoff was \$2.00, if their guess was in a range of $+/-$ 20 Cents of the (rounded) average contribution, their payoff was \$1.00. Otherwise, if their guess was outside the range of $+/-$ 20 Cents, their payoff was \$0.75. This procedure is simple to understand, e.g., compared to a Binarized Scoring Rule (Hossain and Okui 2013), and minimizes background risk, i.e., the risk of receiving a low amount (Selten, Sadrieh, and Abbink 1999). Further, the incentives for making an exact guess were very high (receiving \$2 instead of \$1), so it is implausible that risk-aversion made players deviate from their 'true' guesses.

5.3.4 Post-experimental questionnaire

To investigate the psychological processes behind players' decisions, I elicit their motivations using a post-experimental questionnaire with 7-point scales. First-team players' motivation to set an example for follower-team players was measured as their agreement to the statement, "I wanted that the players in [my follower-team] follow my example.". For follower-team players, I elicited their perception of injunctive norms, i.e., whether they agree that "The players in [the first-team] *should* set an example for [the follower-team players]". I further elicited first-team player's perception of responsibility (e.g., "I did not feel responsible for the players in [my follower-team]). Follower-team players indicated whether first-team players *should* feel responsible for

follower-team players. In addition, I elicited player's motivation for fostering an efficient outcome ("I wanted that everyone gets a higher [payoff]."). Lastly, I also elicited players' interest in finding out how much their co-player contributed to their team's project, and their interest in finding out how much their follower-team contributed, to assess their involvement in the game.

5.4 Hypotheses

First-team players

I consider two principal motives that might explain why people desire to set an example for others, even if there are no monetary incentives for doing so.

Firstly, based on prosocial motivations alone, having social influence on others might lead people to act more cooperatively, thereby aiming to spread cooperation. This might happen if people have a preference for fostering efficient outcomes not only in their own team, but also in their follower-team (Charness and Rabin, 2002).[3] However, two crucial prerequisites for prosocial motivations (e.g., social preferences) to explain people's desire to set an example are that (i) first-team players extend their social considerations to their follower-teams, and (ii) first-team players believe that follower-team players follow their contribution behavior. Note that the first prerequisite is not directly testable, so Hypothesis 1 (below) focuses on the second prerequisite which can be inferred from participants' elicited beliefs.

Hypothesis 1: *Contributions of first-team players who have social influence but could not write a recommendation should increase, but only if they have sufficiently strong beliefs that their contributions lead follower-teams to follow their own contributions.*

Secondly, based on the 'focusing theory' of normative conduct (Krupka and Weber, 2009; Cialdini et al., 1991), people might desire to set an example by making a high contribution, but only when the normative aspects of decision situation become salient. Social norms might draw people's attention to their prosocial identity and self-image as prosocial persons, i.e., people evaluate their own decision relative to the social norm and take their behavior as a signal for their prosocial motivation (Bem, 1972; Bénabou and Tirole, 2011; Leary and Baumeister, 2000). First-team players can voluntarily write a recommendation about what follower-team players *should* contribute, which should make people think about injunctive norms, e.g., that one should make

[3] Also under inequity aversion, for instance, agents perceive the public goods game like a coordination game. Thus, inequity averse first-team players might want to spread equitable cooperative outcomes in their follower-team by making a high contribution in their own team (provided that they have sufficiently optimistic beliefs about their co-players' contribution). Similarly, guilt averse players might want to avoid being responsible for letting down their follower-team's expectation.

a high contribution in their situation. In the *Infl-Rec* treatment, first-team players have social influence on their follower-team players' decisions, so the opportunity to write a recommendation should make it more likely that first-team players perceive a strong injunctive norm to make a high contribution. As a result, a larger number of first-team players might perceive it as socially desirable to set an example, so first-team players' (self-reported) motivation to set an example and their contributions should increase. Conversely, in the treatment *NoInfl-Rec*, the opportunity to write a recommendation should make salient the irrelevance of their decision for others' decisions. In turn, even though players might think about the injunctive norm to make a high contribution, they might perceive that making a high contribution is not necessarily important in their situation, so the own image as a prosocial person is not as much at stake. Hypothesis 2 sums up the expectations due to the salience of injunctive norms and the resulting self-image motives:

Hypothesis 2: *First-team players who have social influence and can write a recommendation should make higher contributions than first-team players who do not have social influence and can write a recommendation, and they should also make higher contributions than first-team players who do not have social influence and cannot write a recommendation.*

Follower-team players

Follower-team player contributions might be affected by first-team players who have social influence in at least two ways. Firstly, if first-team players make a high contribution, follower-team player's beliefs about the average contributions in their follower-team might increase due to descriptive norm effects, i.e., players infer what (they believe) is typically done from the social information (Bernheim, 1994; Cialdini and Goldstein, 2004; Krupka and Weber, 2009; Schultz et al., 2007). Combined with people's tendency to contribute as much as they believe others do, i.e., conditional cooperation (Fischbacher et al., 2001), cooperation might thus increase. Secondly, complementing the first possibility, the social information provided by first-team players might lead to 'focusing' effects (Cialdini et al., 1991; Krupka and Weber, 2009). That is, the social information itself might draw follower-team players' attention to the normative aspects of the decision situation, e.g., to what should be done. For example, Krupka and Weber (2009) found that providing dictators with social information increased the amount shared in a dictator game compared to a situation without social information, even if dictators received information about other dictators' selfish behavior. Similarly, Cialdini et al. (1991) found that a single piece of litter in otherwise clean environment led to less subsequent littering

than an entirely clean environment, suggesting that the single piece of litter made salient the anti-littering norm. Hence, in the present setting, if the mere provision of social information draws follower-team player's attention to the normative aspects of the decision situation, the social information provided should lead follower-teams to more likely follow (normative) high contributions than (counter-normative) low contributions. As a result, contributions might increase in treatments where first-team players have social influence compared to if first-team players do not have social influence, i.e., where decisions cannot be affected by the actual social information received.

Hypothesis 3: *Follower-team players should be more likely to follow high first-team player contributions than low first-team player contributions, if their first-team players have social influence.*

5.5 Results

First-team players

First, I examine whether first-team players expected their contributions to influence follower-team players. If first-team players did not expect any social influence of their own contributions on follower-team players' contributions, then it would be difficult to explain why they nevertheless should have wanted to set an example. In treatments where first-team players had social influence (*Infl-NoRec* and *Infl-Rec*), I elicited first-team players' beliefs about follower-team players' average contributions conditional on three different contribution levels (0, 50, and 100 Cents). Interestingly, on average, first-team players who had social influence expected higher first-team player contributions to positively influence follower-team contributions, regardless of whether first-team players could write a recommendation (see Table 5.1).

Information about first-team player's contribution	0 Cents	50 Cents	100 Cents
First-team players' belief about follower-team players' average contributions in...			
Infl-NoRec	23 Cents	51 Cents	72 Cents
Infl-Rec	27 Cents	47 Cents	69 Cents

Table 5.1: First-team players' beliefs about their social influence on follower-team players' contributions conditional on three elicited contribution levels.

Hence, first-team players believed that higher own contributions lead to higher follower-team

contributions, which is a precondition for prosocial motivations to affect the desire to set an example (Hypothesis 1).

Next, I investigate how first-team players in the different treatments made use of their opportunity to set an example. I do not find support for Hypothesis 1: First-team players who had social influence but could not write a recommendation (*Infl-NoRec*) did not contribute significantly more than first-team players who did not have social influence and could not write a recommendation (*NoInfl-NoRec*) (Permutation test, *Infl-NoRec* vs. *NoInfl-NoRec*, mean 59 Cents vs. mean 56 Cents, Z = -0.57, p = 0.583). First-team players who had social influence but could not write a recommendation (*Infl-NoRec*) also did not contribute significantly more than first-team player who did not have social influence and could write a recommendation (*NoInfl-Rec*) (Permutation test, *Infl-NoRec* vs. *NoInfl-Rec*, 59 Cents vs. 52 Cents, Z = -1.34, p = 0.185). Hence, even though first-team players who had social influence but could not write a recommendation believed that higher contributions increase follower-team player contributions, they did not make significantly higher contributions.

Result 1: *First-team players who had social influence but could not write a recommendation did not contribute significantly more than first-team players who did not have social influence and could not write a recommendation (Infl-NoRec vs. NoInfl-NoRec) or first-team players who did not have social influence and could write a recommendation (Infl-NoRec vs. NoInfl-NoRec).*

By contrast, in support of Hypothesis 2, first-team players who had social influence *and* could write a recommendation (*Infl-Rec*) contributed significantly more than first-team players who did not have social influence and could write a recommendation (Permutation test, *Infl-Rec* vs. *NoInfl-Rec*, mean 65 Cents vs. 52 Cents, Z = -2.41, p = 0.016), and they also contributed more than first-team player who did not have social influence and could not write a recommendation (Permutation test, *Infl-Rec* vs. *NoInfl-NoRec*, mean 65 Cents vs. 56 Cents, Z = -1.67, p = 0.098). First-team players who had social influence and could write a recommendation also contributed more than first-team players who had social influence and could not write a recommendation, but this was not significant (*Infl-Rec* vs. *NoInfl-NoRec*, mean 65 Cents vs. 59 Cents, Z = -1.13, p = 0.268). Figure 5.1 shows the distributions and the medians of first-team players' contribution levels: First-team players who had social influence and could write a recommendation, had a higher median (70) than first-team players in each of the other treatments (50, each). Further, the share of first-team players who contributed 100 Cents (0 Cents) was rather high (low) in the *Infl-Rec* treatment compared to the other treatments. In support of Hypothesis 2, thus, first-

team players only made higher contributions, if they had social influence and the opportunity to write a recommendation made salient injunctive norms.

Result 2: *First-team players who had social influence and could write a recommendation contributed more than first-team players who did not have social influence and could write a recommendation (Infl-Rec vs. NoInfl-Rec) or could not write a recommendation (Infl-Rec vs. NoInfl-NoRec).*

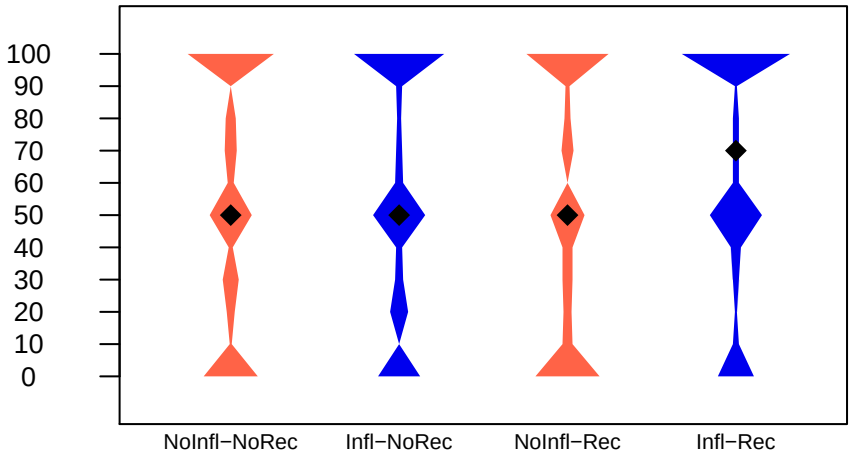


Figure 5.1: Violin plots showing the rotated distributions and medians of first-team player's contributions for each treatment.

Next, I investigate first-team players' motivations as stated in the post-experimental questionnaire. In doing so, I can find out how first-team players who had influence and could write a recommendation (*Infl-Rec*) perceived their decision situation and test how their perception differed from (i) the perception of first-team players who did not have social influence (*NoInfl-NoRec*, *NoInfl-Rec*), and (ii) the perception of first-team players who also had social influence but could not write a recommendation. Potentially, players who had social influence and could write a recommendation about what follower-team players should contribute were more likely to focus on injunctive norms and thus reported a stronger motivation to set an example (supporting Hypothesis 2). In line with this reasoning, first-team players who had social influence and could write a recommendation reported a higher motivation to set an example than first-team players who did not have social influence and could write a recommendation (*Infl-Rec* vs. *NoInfl-Rec*, 5.70 vs. 4.51, Z = -4.16, $p < .001$) or could not write a recommendation (*Infl-Rec* vs. *NoInfl-NoRec*, 7-point-scale, 5.70 vs. 4.73, Z =

-3.52, $p < .001$). Importantly, first-team players who had social influence and could write a recommendation also reported a higher motivation to set an example than first-team players who had social influence and could not write a recommendation (*Infl-Rec* vs. *Infl-NoRec*, 7-point-scale, 5.70 vs. 4.96, Z = -2.74, $p = .006$).[4]

With regard to the other questionnaire variables, first-team players who had social influence and could write a recommendation reported significantly stronger feelings of responsibility and a stronger interest in their follower-team players' behavior compared to first-team players who did not have social influence and could write a recommendation (*NoInfl-Rec*) or could not write a recommendation (*NoInfl-Rec*) (Permutation tests, all p-values $< .05$). However, the first-team players who had social influence and could write a recommendation did not report stronger feelings of responsibility for their follower-team players, a higher interest in their follower-team players' behavior, or a stronger desire to foster efficient outcomes compared to first-team players who had social influence but could not write a recommendation (*Infl-Rec* vs. *Infl-NoRec*, all p-values $> .10$). Thus, while other questionnaire measures also affected the perceptions of first-team players who had social influence and could write a recommendation (*Infl-Rec*), the particularly high contributions in this treatment are best attributed to first-team players' increased motivation to set an example.

Result 3: *If first-team players had social influence and could write a recommendation, they reported a higher motivation to set an example compared to first-team players who did not have social influence (Infl-Rec vs. NoInfl-NoRec, Infl-Rec vs. NoInfl-Rec) and compared to first-team players who had social influence but could not write a recommendation (Infl-Rec vs. Infl-NoRec).*

Norm perceptions

Next, I report evidence about player's norm perceptions to be able to better interpret the follower-team player's behavior. Interestingly, the maximum contribution of 100 Cents was the most frequent recommendation about what follower-team players should do (59 of 103 recommendations in *Infl-Rec*, 40 out of 72 recommendations in *NoInfl-Rec*). Hence, the maximum contribution was most likely also what participants considered the injunctive norm. To infer whether participants expect first-team players to set an example from a normative

[4] A mediation analysis suggests that the stronger motivation to set an example for follower-team players also led to higher contributions as compared to the other treatments: Controlling for all other questionnaire measures, assignment to the treatment *Infl-Rec* predicted a higher motivation to set an example than in other treatments ($b_{Infl-Rec} = 0.499$, SE = 0.178, $p = .005$), higher levels of which in turn predicted higher contributions ($b_{\text{example}} = 4.301$, SE = 0.880, $p < .001$ (indirect effect, E[contribution | example(Infl-Rec = 1)] – E[contribution | example(Infl-Rec = 0)] = 2.133, CI95 [0.56,4.08], $p = .006$).

perspective, I can look at follower-team players' self-reported agreement that first-team players should set an example in the post-experimental questionnaire. Notably, follower-team players' perception that first-team players should set an example was near maximum agreement in treatments where first-team players had social influence, and higher than in treatments where first-team players did not have social influence (7-points-scale, means: *Infl-Rec*, 5.64, *Infl-NoRec*, 5.84, *NoInfl-Rec*, 3.97, *NoInfl-NoRec*, 4.00, all pairwise comparisons between *Infl* and NoInfl treatments, $p < .001$).[5]

Result 4: *First-team players' recommendations suggest that they perceive an injunctive norm to contribute the full amount of 100 Cents. Alike, follower-team players reported strong agreement that first-team players should set an example, if first-team players had social influence (Infl-Rec and Infl-NoRec).*

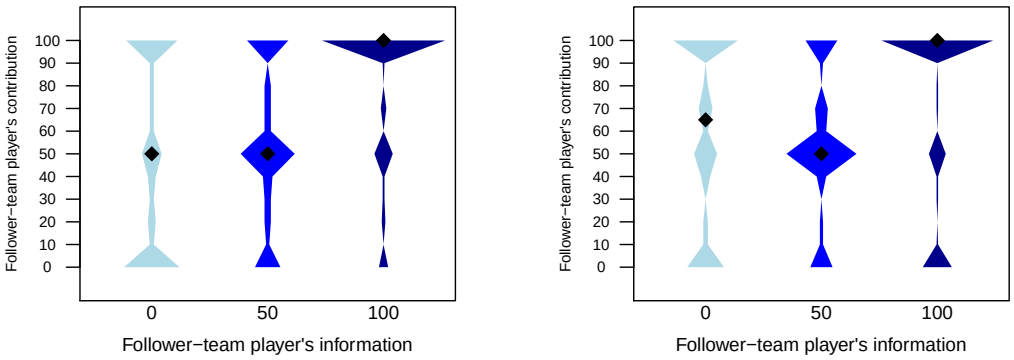
Follower-team players

In the following, I average across the contributions of each follower-team (the level of independent observation) to analyze the social influence of first-team players' contributions on follower-team players' contributions. If first-team players had social influence, descriptive norm effects would suggest that higher (lower) first-team player contributions should translate into higher (lower) follower-team player contributions. Looking at mere correlations, the first-team players' and their follower-team's average contributions were significantly correlated in both treatments where first-team players had social influence (*Infl-NoRec*, $\rho = .58$, $p < .001$, *Infl-Rec*, Sparman's $\rho = .30$, $p < .001$; see also Appendix 5.A for regressions). To see how follower-team players expected the social information to affect beliefs, I elicited follower-team player's beliefs about their co-player's contributions levels conditional on different contribution levels by first-team players (0, 50, and 100). In line with descriptive norm effects, on average, follower-team players expected higher first-team player contributions to coincide with higher follower-team player contributions (see Table 5.2).

[5]Note that first-team players were not asked whether they themselves thought that they should set an example, since such an elicitation might confound their reports about their motivation to set an example.

Information about first-team player's contribution	0 Cents	50 Cents	100 Cents
Follower-team players' belief about follower-team players' average contributions in...			
Infl-NoRec	30 Cents	53 Cents	77 Cents
Infl-Rec	27 Cents	47 Cents	70 Cents

Table 5.2: Follower-team players' beliefs about the social influence of first-team players' contributions on follower-team players' contributions conditional on three elicited contribution levels.



(a) *Infl-NoRec.* **(b)** *Infl-Rec.*

Figure 5.2: Violin plots showing the rotated distributions and medians of follower-team player's contributions for different information about contribution levels of first-team players. Only the contribution levels of 0, 50 and 100 are shown as these make up about 75% of the first-team player contributions, whereas the remainder of the data is scattered evenly across the remaining contribution levels.

As can be seen in Figure 5.2, however, follower-team players' behavior differed markedly from mere descriptive norm effects. The figure shows how the distribution of follower-team player contributions varies with the contribution information received in treatments with social influence (*Infl-NoRec* and *Infl-Rec*). Notably, when receiving information about a zero contribution of their first-team player, the distribution of follower-team player contributions in the treatments with social influence (*Infl-NoRec* and *Infl-Rec*) was not skewed towards zero contributions and it did not differ significantly to the distribution of follower-team contributions in treatments without social influence (*NoInfl-NoRec* and *NoInfl-Rec*) (*Infl-NoRec*, information: 0 Cents, mean contribution 48 Cents; *NoInfl-NoRec*, overall mean contribution 57 Cents; *Infl-Rec*, information: 0 Cents, mean contribution 59 Cents; *NoInfl-Rec*, overall mean contribution 57 Cents, Kolomogorov-Smirnov tests comparing the distributions, all p-values $> .378$).[6] Hence, follower-team players hesitated to follow low contributions. This

[6]The high contributions of follower-teams who receive information about a 0 Cents contribution in *Infl-Rec* can *in part* be attributed to recommendations which are inconsistent with behavior: Eight first-team players

is indicative of Hypothesis 3, i.e., the social information 'focused' participants on the normative aspects of the decision situation, so they refrain from imitating 'counter-normative' behavior. Also in line with 'focusing' effects, when receiving information about a 100 Cents contribution of their first-team player, the distribution of follower-team players' contributions is highly skewed, with the vast majority of follower-team players contributing 100 Cents. In this case, the distribution of follower-team player contributions in treatments with social influence (*Infl-NoRec* and *Infl-Rec*) differs substantially from the distribution of follower-team player contributions in treatments without social influence (*NoInfl-NoRec* and *NoInfl-Rec*) (*Infl-NoRec*, information: 100 Cents, mean contribution 85 Cents; *NoInfl-NoRec*, overall mean contribution 57 Cents; *Infl-Rec*, information: 100 Cents, mean contribution 75 Cents; *NoInfl-Rec*, overall mean contribution 57 Cents; Kolomogorov-Smirnov tests comparing the distributions, all p-values $< .004$). Result 5 sums up the findings supporting Hypothesis 3.[7]

Result 5: *If first-team players had influence (Infl-Rec and Infl-NoRec), follower-team players were more likely to follow high contributions than they were to follow low contributions.*

Next, I compare follower-team contributions across different treatments to see whether first-team players who had social influence positively influenced follower-team player contributions in the aggregate. Follower-team players whose first-team players had social influence but could not write a recommendation made higher contributions than follower-team players whose first-team players did not have social influence and could not write a recommendation (Permutation test, *Infl-NoRec* vs. *NoInfl-NoRec*, mean 64 Cents vs. 57 Cents, Z = -2.01, p = .045, see also Figure 5.3) and they also made higher contributions compared to follower-team players whose first-team players did not have social influence and could write a recommendation (*Infl-NoRec* vs. *NoInfl-Rec*, 64 Cents vs. 57 Cents, Z = -1.88, p = .061). Hence, in *Infl-NoRec*, follower-team players' contributions increased even though the contributions of their first-team players who provided the information did not increase relative to treatments where first-team players did not have social influence (cf. Result 1 and Figure 5.1). The increase might thus result from 'focusing

contributed 0 Cents, but recommended a 100 Cents contribution; resulting in average follower-team contributions of 67 Cents. Yet, note that the remaining follower-team players that received information about a 0 Cents contribution in *Infl-Rec* also contributed 52 Cents on average and contributions were similarly high (48 Cents) in *Infl-NoRec*, i.e., where follower-team players do not receive any (inconsistent) recommendations.

[7]Regression analyses in the Appendix 5.A also support Result 5. The share of follower-team players that contributed exactly the same amount which their first-team player contributed, i.e., who "follow" their first-team players, was significantly larger in treatments where first-team players had social influence than in treatments where they did not have social influence (e.g., comparing *Infl-Rec* vs. *NoInfl-Rec*, $b_{information*NoInfl-Rec}$ = -0.026, SE = .010, p = .012, Infl-Rec vs. *NoInfl-NoRec*, $b_{information*NoInfl-NoRec}$ = -0.032, SE = .010, p = .002). Note that in the *NoInfl-Rec* and *NoInfl-NoRec* treatments, follower-team contributions can coincide with their first-team player contributions only by chance.

effects' (Hypothesis 3), i.e., the social information made salient the normative aspects of the decision situation.

Follower-team players whose first-team players had social influence and could write a recommendation also made higher contributions than follower-team players whose first-team players did not have social influence and could write a recommendation (*Infl-Rec* vs. *NoInfl-Rec*, 65 Cents vs. 57 Cents, Z = -2.13, p = .034), and they also made higher contributions than follower-team players whose first-team players did not have social influence and could not write a recommendation (*Infl-Rec* vs. *NoInfl-NoRec*, 65 Cents vs. 57 Cents, Z = -2.26, p = .024, see also Figure 5.3). While follower-team players' contributions were relatively high in *Infl-Rec*, contributions might not have increased further, e.g., compared to *Infl-NoRec*, due to ceiling effects, or because some of the first-team players' recommendations contain additional information. Result 6 sums up the findings on aggregate follower-team contributions.

Result 6: *Follower-team players' contributions increased significantly if first-team players had social influence (Infl-Rec and Infl-NoRec) as compared to if they did not have social influence (NoInfl-Rec and NoInfl-NoRec).*

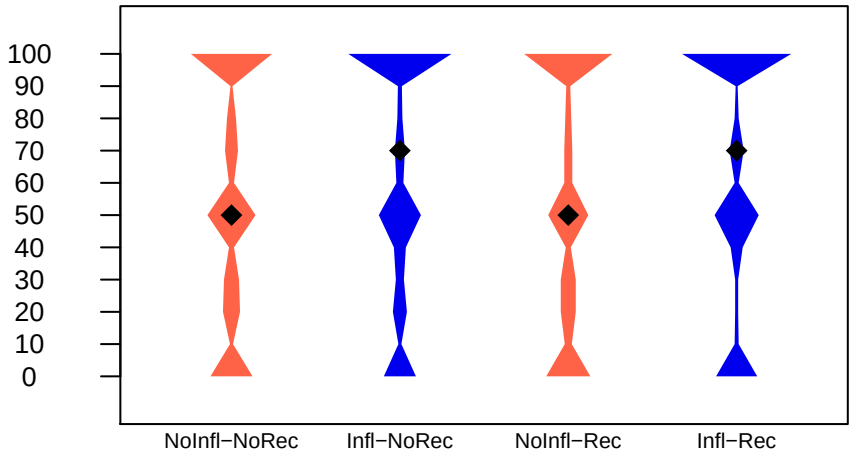


Figure 5.3: Violin plots showing the rotated distributions and medians of follower-team player's contributions for each treatment.

5.6 Discussion and Conclusion

This paper investigates whether or not people desire to set an example for others, if there are no monetary incentives for doing so. Players of one team ('first-team' players) played the

public goods game and the decisions in their team could affect the social information available to others who independently played their own public goods game in so-called 'follower-teams'. Each first-team player either had social influence over his or her follower-team (treatments *Infl-NoRec* and *Infl-Rec*) or did not have social influence (treatments *NoInfl-NoRec* and *NoInfl-Rec*). Further, to vary the salience of injunctive norms and thus of people's self-image as a prosocial person, first-team players either could additionally write a recommendation about what follower-team players *should* do (treatments *Infl-Rec* and *NoInfl-Rec*) or they could not do so (treatments *Infl-NoRec* and *NoInfl-NoRec*). If first-team players had social influence but could not write a recommendation (*Infl-NoRec*), prosocial motivations alone might explain higher contributions. Yet, even though first-team players in this treatment on average believed to have social influence, their contributions did not increase significantly. Only when first-team players had social influence and could write a recommendation (*Infl-Rec*), they contributed significantly more than first-team players who did not have social influence on their follower-team players' contributions and could (or could not) write a recommendation (*NoInfl-Rec* and *NoInfl-NoRec*). This suggests that merely having social influence did not suffice to increase contributions. Instead, only if people's attention was drawn to injunctive norms (Cialdini et al., 1991), thereby activating people's desire for prosocial identity and positive self-image (Bem, 1972; Bénabou and Tirole, 2011; Leary and Baumeister, 2000), people set an example for others by making higher contributions.

Further analyses help substantiate the suspected psychological mechanism. While first-team players who had social influence and could write a recommendation (*Infl-Rec*) did not report a stronger motivation for fostering an efficient outcome than first-team players who had social influence but could not write a recommendation, they did report a stronger motivation to set an example for their follower-team players. Most likely, the increased motivation to set an example thus arose because first-team players were made aware of the injunctive norms, so their self-image as a prosocial person was at stake. Hence, this study highlights the importance of people's desire for positive self-image as a prosocial person for explaining why people set an example for others.[8]

In addition, this paper documents interesting findings regarding the impact of social information on 'conditional cooperation' (Kelley and Stahelski, 1970; Fischbacher et al., 2001): The availability of social information in treatments where first-team players had social influence led to increased contributions in follower-teams compared to treatments where

[8]Note that also a desire to act consistently with their own recommendation, e.g., to avoid cognitive dissonance (Konow, 2000), or a desire for social approval from follower-team players might have motivated contributions. Yet, these explanations ultimately also entail self-image concerns that are closely related to normative expectations.

first-team players did not have social influence. The follower-team players' high contributions can be attributed to their inclination to follow high but not low first-team players' contributions. The follower-team players thus did not only follow the contribution information they received, as mere descriptive norm effects would suggest (Bernheim, 1994; Cialdini and Goldstein, 2004; Krupka and Weber, 2009). Rather, in addition, there were 'focusing' effects, i.e., the social information made the normative aspects of the decision situation more salient, so people hesitated to follow counter-normative contribution examples.[9] This explanation is also consistent with the findings of one of the most widely cited dictator-game study on social influence (Krupka and Weber, 2009). The authors found increases in dictator game giving in treatments with social information compared to a control without social information, where participants received information about selfish behavior by previous dictators. Similarly, a recent field experiment in a firm found increases in donations when employees were informed that a share of only about 5% of them typically donate compared to a control without social information (Sanders, 2017).

The finding of rather high contributions in follower-teams are intriguing in the view of 'selfish-biases' in conditional cooperation. Fischbacher and Gächter (2010) found that people cooperate conditionally, but contribute slightly less than others. Similarly, in the leading-by-example setting, 'followers' of the leader typically reciprocate only partly by making lower contributions than their leader. An explanation for this apparent divergence is that follower-team players evaluated the social information from a normative perspective, whereas in the other settings free-riding incentives were more salient. Further, compared to Fischbacher and Gächter (2010), the marginal-per-capita-return to cooperation (MPCR) in the present study was relatively high (0.75 compared to 0.4). Perhaps, thus, biases in conditional cooperation might also favor cooperation, if cooperation appears sufficiently attractive.

From a policy perspective, the findings of the present study imply that – to appeal to people's desire to set an example and increase cooperation – merely making their contributions public does not suffice. Instead, in addition, injunctive norms need to be made salient, e.g., by having people make personal recommendations about what others should do. Further, the evidence that normative behaviors (high contributions) were easier to spread than counter-normative behaviors (low contributions) might explain why norms – be they 'good' or 'bad' – have proven 'sticky' and difficult to change in the past (as is the case for corruption, e.g., (Becker et al., 2016)).

[9] If first-team players had social influence, the maximum contribution was the first-team player's most frequent recommendation, and follower-team players strongly agreed that first-team players should set an example.

Appendices

5.A Regressions analyzing social influence

To analyze more closely how the social information affected the follower-team players' behavior, I analyze (i) their contributions, and (ii) the share of follower-team players who contribute exactly the same amount which their first-team player contributed, i.e., the share of 'followers'. Notably, an increase in the share of 'followers' might also occur by chance due to the higher frequency of high compared to low contribution. To account for this possibility, I control for social information in treatments where first-team players did not have social influence, so the social information was not available when making the decision and any correlation can occur only by chance. To this end, I use a linear regression that aggregates follower-team player's contributions across teams (see Table 5.3). As dependent variables, I use the average contribution of players of the same follower-team (Model 1), and the share of 'followers' per team, that is, the share of follower-team players of the same team who contribute exactly the amount which their first-team player did, i.e., 0%, 50%, or 100% (Model 2). As independent variables, I control for the treatment dummies and a continuous variable info with the information about contribution level of the follower-team's first-team player (which is divided by 10 to obtain the impact of a 10 Cents increase). I interact treatment dummies and the information about the first-team player's contribution level to detect between-treatment differences in behavior. I use the treatment *Infl-Rec* as a reference category. Model 1 confirms the expected impact of social information on contributions in treatments where first-team players had social influence, e.g., a 10 Cents increase in first-team player's contributions led to roughly a 2 Cents increase in contributions in follower-teams (b_{Info} = 2.09, SE = 0.62, $p < .001$). In Model 2, I find that for each 10 Cents-increase in the information received by a first-team player, there is roughly a 5% increase in the share of followers if first-team players had social influence and could write a recommendation (*Infl-Rec*, b_{Info} = 0.049, SE = .008, $p < .001$), which is significantly larger than the increase that occurs by chance due to the higher frequency of high compared to low contributions in treatments without social influence (e.g., comparing *Infl-Rec* vs. *NoInfl-Rec*, b_{Info}*NoInfl-Rec = -0.026, SE = .010, p = .012, *Infl-Rec* vs. *NoInfl-NoRec*, b_{Info}**NoInfl-NoRec* = -0.032, SE = .010, p = .002). Estimates for *Infl-NoRec* are very similar to that of *Infl-Rec* (estimated from Model 2, *Infl-NoRec*, b_{Info} = 0.047, SE = .008, $p < .001$).

	Contributions (1)	Followers (2)
Info	2.09***	0.05***
	(0.62)	(0.01)
Infl-NoRec	−12.51*	0.04
	(6.31)	(0.08)
NoInfl-NoRec	8.47	−0.05
	(6.15)	(0.07)
NoInfl-Rec	7.78	−0.12
	(5.95)	(0.07)
Info*Infl-NoRec	2.18*	−0.00
	(0.87)	(0.01)
Info*NoInfl-NoRec	−2.57**	−0.03**
	(0.86)	(0.01)
Info*NoInfl-Rec	−2.42**	−0.03*
	(0.84)	(0.01)
(Intercept)	51.02***	0.17**
	(4.63)	(0.06)
R^2	0.13	0.30
Adj. R^2	0.12	0.29
Num. obs.	480	480

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\circ}p < 0.1$

Table 5.3: Linear regressions showing the impact of social information on conditional cooperation in the follower-teams (average contributions in Model1, share of followers in Model 2).

5.B Instructions

All treatments:

Please **only continue** to work on this HIT if you can work without interruption for the next **15 minutes**.

You will be able to enter your Worker ID only once.

Continue

Please solve the recognition task below.

five plus one is

Please enter your Worker ID:

ID

Your Worker ID is needed to ensure you get your payoff.
You can only participate once.
If you don't know your Worker ID you can find it out on the following page (opens in a new tab):
https://www.mturk.com/mturk/dashboard

Back Next

Participation conditions

To get your HIT accepted, please answer all control questions correctly. Read carefully!

Your **acceptance payment: $0.75**.

You can earn a ***significant*** **bonus payment** of at least $0.75.

Your **bonus payment** depends on your decisions and the decisions of other MTurk workers.

Your total payoff = acceptance payment + bonus payment

Procedure

We recruit more than 300 MTurk workers (including you). Other MTurk workers either play before or after you. When all MTurk workers have played, you will be randomly matched to other MTurk workers and you will receive your bonus within **three work days**.

Anonymity

Your MTurk Worker ID and total payoffs are confidential: only you and the creator of the survey will know the amount of money you make.

Next

Please note:

You will see two decision situations. Only one of the two decision situations will be randomly selected for your bonus payment.

You will first read the "Instructions for the Contributions Game", which will be relevant for both decision situations.

Back Next

You will need to answer control questions **from memory**.

You will not be able to review the instructions more than once.

Instructions: The Contribution Game

We will randomly match **you and another MTurk worker** into a pair.

In your pair, you and your co-player **choose independently** how many Cents to contribute to your pair´s project.

You and your co-player both face the **identical decision situation**:

- Each player receives **100 Cents** ($1.00).
- Of these 100 Cents, each player can **contribute** between 0 Cents and 100 Cents **to your pair´s project**.
- Each Cent that a player does **not contribute** to your pair´s project, he/she will keep.
- **Each player profits from each Cent contributed** to your pair´s project.
- More precisely: Each Cent contributed to your pair´s project is **multiplied by 1.5 and redistributed equally** among both players. This means that for **each 10 Cents contributed** to your pair´s project, 15 Cents are redistributed equally among both players, so **each player receives 7.5 Cents**.

Your bonus in the Contribution Game:

Cents kept by you+
7.5 Cents for each 10 Cents which you contributed +
7.5 Cents for each 10 Cents which your co-player contributed

Next

Please answer the question below.

You can **test** how the calculations work by trying different contributions for both you and your co-player:

Your contribution:

0 Cents — 100 Cents

50 Cents

Your co-player´s contribution:

0 Cents — 100 Cents

50 Cents

Your bonus:	125 Cents
Your co-player´s bonus:	125 Cents

Question 1

To get your HIT accepted, please mark the correct statement(s)!

For a correct answer you may need to select more than one statement! You have **three trials in total** to mark the correct statement(s).

- [ ] Every Cent which a player does not contribute to your pair´s project, he/she will keep.
- [ ] If you contribute 0 Cents and your co-player contributes 100 Cents, your bonus is 75 Cents and your co-player's bonus is 175 Cents.
- [ ] If you and your co-player both contribute 100 Cents, your bonus and your co-player's bonus is 150 Cent.
- [ ] If you and your co-player both contribute 0 Cents, your bonus and your co-player's bonus is 100 Cents.
- [ ] If you contribute 0 Cents and your co-player contributes 100 Cents, your bonus is 175 Cents and your co-player's bonus is 75 Cents.
- [ ] Every Cent that is contributed to your project will be multiplied by 1.4.

Review Instructions (only once) Submit Answer

Infl-Rec (Infl-NoRec without hints on recommendation):

You will need to answer control questions **from memory**.

You will not be able to review the instructions more than once.

Instructions: Decision Situation One

You and five other MTurk workers will be randomly matched into one of three pairs: Pair A, Pair B, or Pair C.
Each pair will play its own Contribution Game **separately.**

First, the players in **Pair A** make their **contributions** to their pair´s project.

Each player in Pair A provides the following **information** for players of a different pair **who have not yet made their contributions**
(Pair B or Pair C):

- **information on his/her contribution** *and*
- **a recommendation about what the players should contribute** (voluntary).

One player in Pair A provides information for Pair B. The other player in Pair A provides information for Pair C.

Second, the players in **Pair B and Pair C** each **receive the information** on the contribution and the recommendation of their player in Pair A.

At last, the players in Pair B and Pair C each make their **contributions** to their pair´s project.

NOTE: When the players in Pair B or Pair C **receive the information** on the contribution and the recommendation, they have **not yet made their contributions** to their pair´s project.

Your bonus in Decision Situation One depends only on how much you and your co-player in your pair contribute to your pair´s project.

Next

Question 2

To get your HIT accepted, please mark the correct statements!

You have **three trials in total** to mark the correct statement(s).

The players in Pair A play...

- ○ in one common Contribution Game jointly with the other pairs.
- ○ their own Contribution Game separately from other pairs.

One player in Pair A provides information for Pair B ...

- ○ and this player also provides information for Pair C.
- ○ and the other player in Pair A provides information for Pair C.

Who receives information on the contribution from a player in Pair A...

- ○ his/her co-player in Pair A.
- ○ the players in Pair B or Pair C.

When the players in Pair B receive the information of a player in Pair A...

- ○ they **cannot change** their contributions anymore.
- ○ they have **not yet** made their contributions to their pair´s project.

Review Instructions (only once) Submit Answer

You have been randomly matched into **Pair A**.

Back Next

Decision Screen (Pair A)

You play the Contribution Game with your co-player in Pair A.

Please make your **CONTRIBUTION** to your pair´s project.

0 Cents — 100 Cents

You provide **information** for the two players in **Pair B**.

The players in Pair B **play** their own Contribution Game **separately** from your pair.

If you want, you can make a **recommendation about what the two players in Pair B should contribute**:

Note: The players in Pair B receive your recommendation **before** they contribute.

Both players in Pair B receive information on your contribution and your recommendation before they make their contributions:

Decision Screen (PAIR B)

Information

Your player in Pair A contributed Cents.

When you receive your bonus, you will be **informed** about

- the contribution of your co-player in Pair A,
- and the contributions of both players in Pair B.

Back | Make Your Contribution

NoInfl-Rec (NoInfl-NoRec without hints on comment):

You will need to answer control questions **from memory**.

You will not be able to review the instructions more than once.

Instructions: Decision Situation One

You and five other MTurk workers will be randomly matched into one of three pairs: Pair A, Pair B, or Pair C.
Each pair will play its own Contribution Game **separately.**

First, the players in **Pair A** make their **contributions** to their pair´s project.

Each player in Pair A provides the **following information** for players of a different pair **who have already finished their game** (Pair B or Pair C):

- **information on his/her contribution** *and*
- **a comment about what the players should have contributed** (voluntary) .

One player in Pair A provides information for Pair B. The other player in Pair A provides information for Pair C.

Second, the players in **Pair B and Pair C** each make their **contributions** to their pair´s project.

At last, the players in Pair B and Pair C each **receive the information** on the contribution and the comment of their player in Pair A.

NOTE: When the players in Pair B or Pair C **receive the information** on the contribution and the comment, they **cannot change their contributions** anymore.

Your bonus in Decision Situation One depends only on how much you and your co-player in your pair contribute to your pair´s project.

Next

Question 2

To get your HIT accepted, please mark the correct statements!

You have **three trials in total** to mark the correct statement(s).

The players in Pair A play...

- ○ their own Contribution Game separately from other pairs.
- ○ in one common Contribution Game jointly with the other pairs.

Who receives information on the contribution from a player in Pair A...

- ○ the players in Pair B or Pair C.
- ○ his/her co-player in Pair A.

One player in Pair A provides information for Pair B ...

- ○ and this player also provides information for Pair C.
- ○ and the other player in Pair A provides information for Pair C.

When the players in Pair B receive the information of a player in Pair A...

- ○ they **cannot change** their contributions anymore.
- ○ they have **not yet** made their contributions to their pair´s project.

Review Instructions (only once) Submit Answer

Decision Screen (Pair A)

You play the Contribution Game with your co-player in Pair A.

Please make your contribution to your pair´s project.

0 Cents ——— 100 Cents

You provide information for the two players in **Pair B**.

The players in Pair B play their own Contribution Game separately from your pair.

If you want, you can make a **comment about what the two players in Pair B should have contributed**:

Note: The players in Pair B have already finished and cannot change their contributions anymore.

Both players in Pair B receive information on your contribution and your comment after they have finished their game:

Information

Your player in Pair A contributed Cents.

When you receive your bonus, you will be **informed** about

- the contribution of your co-player in Pair A,
- and the contributions of both players in Pair B.

Back | Make Your Contribution

Infl-Rec and Infl-NoRec (adapted in the NoInfl treatments)

Decision Situation Two

You need to **guess as accurately as possible** the contributions of other MTurk workers in Decision Situation One.

There are **two tasks**: the **PAIR A TASK** and the **PAIR B/C TASK**.

If Decision Situation Two is selected for your bonus payment, **one of the two tasks** will be randomly selected for your bonus payment.

Back Next

PAIR A TASK

You guess the **AVERAGE CONTRIBUTION** of players in **Pair A** in Decision Situation One.

Your **GUESS is compared to the AVERAGE CONTRIBUTION** of actual players in Pairs A.

Note: The AVERAGE CONTRIBUTION is rounded to the nearest 10 Cents.

Your bonus in the PAIR A TASK:

- We add **$2.00 to your bonus**, if your GUESS is
  - **equal** to the AVERAGE CONTRIBUTION.
- We add **$1.00 to your bonus**, if your GUESS is
  - at most **20 Cents above** the AVERAGE CONTRIBUTION,
  - or at most **20 Cents below** the AVERAGE CONTRIBUTION.
- Otherwise, we add $0.75 to your bonus.

Note: The AVERAGE CONTRIBUTION is rounded to the nearest 10 Cents.

Your GUESS: What is the **AVERAGE CONTRIBUTION** of players in **Pairs A**?

0 Cent 100 Cent

Back Submit Your GUESS

PAIR B/C TASK

You **GUESS** the AVERAGE CONTRIBUTION of players

- in **Pairs B and Pairs C** in Decision Situation One
- who received **identical information on the contribution** of their player in Pair A.

Recall: When the **players in Pair B or Pair C** receive the information, they **have not yet made their contributions** to their pair´s project.

Review Decision Situation

Preview:

Your GUESS: What is the **AVERAGE CONTRIBUTION of players** in Pairs B and Pairs C **who received the information**...?

One of your GUESSES in the PAIR B/PAIR C TASK is randomly selected for your bonus payment.

The selected **GUESS is compared to the AVERAGE CONTRIBUTION** of the players in Pairs B and Pairs C who received identical information.

Your bonus in the PAIR B/C TASK:

- We add **$2.00 to your bonus**, if your GUESS is
  - **equal** to the AVERAGE CONTRIBUTION.
- We add **$1.00 to your bonus**, if your GUESS is
  - at most **20 Cents above** the AVERAGE CONTRIBUTION,
  - or at most **20 Cents below** the AVERAGE CONTRIBUTION.
- Otherwise, we add $0.75 to your bonus.

Note: The AVERAGE CONTRIBUTION is rounded to the nearest 10 Cents.

Question 3

To get your HIT accepted, please mark the correct statement(s)!

A correct answer may require selecting more than one statement!
You have **three trials in total** to mark the correct statement(s).

- [ ] You GUESS the AVERAGE CONTRIBUTION of players that received identical information on the contribution of their player in Pair A.
- [ ] We add $1.00 to your bonus, if your GUESS is 21 Cents below the AVERAGE CONTRIBUTION.
- [ ] We add $2.00 to your bonus, if your GUESS equals the AVERAGE CONTRIBUTION.
- [ ] Precisely one of your GUESSES in the PAIR B/C TASK is randomly selected for your bonus payment.

Back | Answer Question

Guess Pair B and Pair C

Recall: When the **players in Pair B or Pair C** receive the information, they **have not yet made their contributions** to their pair´s project.

Your GUESS: What is the **AVERAGE CONTRIBUTION of players** in Pairs B and Pairs C **who received the information**...?

	0 Cent	100 Cent
... "Your player in Pair A contributed 0 Cents."		
... "Your player in Pair A contributed 50 Cents."		
... "Your player in Pair A contributed 100 Cents."		

Back | Submit Your all Your GUESSES

To complete, please indicate to what extent you agree with the following five statements...

When deciding about my contribution, ...	do not agree — strongly agree
I wanted that players in Pair B follow my example.	
I did not feel responsible for the decisions of players in Pair B.	
I wanted that everyone earns a higher bonus.	

When receiving the bonus, ...	do not agree — strongly agree
I am keen on finding out about the decision of my co-player.	
I am keen on finding out about the decisions of players in Pair B.	

Please complete...

Your gender (male, female, other)	
Your age (in years)	
Your understanding of 'average': What is the average of 2, 2, 5?	

Back | Proceed to Payoff and Completion Code

Total Payoff

Your acceptance payoff is $0.75.

Your **bonus payment** will be at least $0.75.

Your bonus payment will depend on the other players' decisions and the decision situation selected for your bonus payment.

We will inform you about this in your bonus message.

Next

Bibliography

Abbink, K., Brandts, J., Herrmann, B., Orzen, H., 2012. Parochial altruism in inter-group conflicts. Economics Letters 117 (June), 45–48.

Abbink, K., Hennig-Schmidt, H., 2006. Neutral versus loaded instructions in a bribery experiment. Experimental Economics 9 (2), 103–121.

Adena, M., Huck, S., 2016. Online fundraising, self-deception, and the long-term impact of ask avoidance (April).

Allcott, H., 2011. Social norms and energy conservation. Journal of Public Economics 95 (9-10), 1082–1095.

Almås, I., Cappelen, A. W., Salvanes, K. G., Sørensen, E. Ø., Tungodden, B., dec 2015. Willingness to Compete: Family Matters. Management Science 62 (8), 2149–2162.

Almås, I., Cappelen, A. W., Sørensen, E. Ø., Tungodden, B., 2010. Fairness and the development of inequality acceptance. Science (New York, N.Y.) 328 (5982), 1176–1178.

Altermatt, E. R., Pomerantz, E. M., 2005. The implications of having high-achieving versus low-achieving friends: A longitudinal analysis. Social Development 14 (1), 61–81.

Ames, D. R., 2004. Strategies for Social Inference: A Similarity Contingency Model of Projection and Stereotyping in Attribute Prevalence Estimates. Journal of Personality and Social Psychology 87 (5), 573–585.

Amir, O., Rand, D. G., Gal, Y. K., 2012. Economic games on the internet: The effect of $1 stakes. PLoS ONE 7 (2), 1–4.

Anderson, L. R., Mellor, J. M., Milyo, J., 2008. Inequality and public good provision: An experimental analysis. The Journal of Socio-Economics 37 (3), 1010–1028.

Andreoni, J., Bernheim, B. D., 2009. Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects. Econometrica 77 (5), 1607–1636.

Andreoni, J., Rao, J. M., Trachtman, H., 2017. Avoiding the ask: a field experiment on altruism, empathy, and charitable giving. Journal of Political Economy 125 (3), 625–653.

Arbak, E., Villeval, M. C., 2013. Voluntary leadership: Motivation and influence. Social Choice and Welfare 40 (3), 635–662.

Archibald, S. J., Kerns, K. a., 1999. Identification and Description of New Tests of Executive Functioning in Children. Child Neuropsychology 5 (2), 115–129.

Ariely, D., Bracha, A., Meier, S., 2009. Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. American Economic Review 99 (1), 544–555.

Babcock, L., Wang, X., Loewenstein, G., 1996. Choosing the Wrong Pond: Social Comparisons in Negotiations That Reflect a Self-Serving Bias. Quarterly Journal of Economics 111 (1), 1–19.

Bakshy, E., Messing, S., Adamic, L. A., 2015. Exposure to ideologically diverse news and opinion on Facebook. Science 348 (6239), 1130–1132.

Balliet, D., Wu, J., De Dreu, C. K. W., 2014. Ingroup Favoritism in Cooperation: A Meta-Analysis. Psychological Bulletin 140 (6), 1556–1581.

Banerjee, R., 2000. The development of an understanding of modesty. British Journal of Developmental Psychology 18, 499–517.

Banerjee, R., Bennett, M., Luke, N., 2010. Children's reasoning about the self-presentational consequences of apologies and excuses following rule violations. British Journal of Developmental Psychology 28, 799–815.

Barclay, P., Willer, R., 2007. Partner choice creates competitive altruism in humans. Proceedings of the Royal Society B: Biological Sciences 274 (1610), 749–753.

Bardsley, N., 2008. Dictator game giving: altruism or artefact? Experimental Economics 11 (2), 122–133.

Bartling, B., Engl, F., Weber, R. A., 2014. Does willful ignorance deflect punishment? - An experimental study. European Economic Review 70, 512–524.

Bartling, B., Fischbacher, U., Schudy, S., 2015. Pivotality and responsibility attribution in sequential voting. Journal of Public Economics 128, 133–139.

Bartling, B. B., Fischbacher, U., 2012. Shifting the Blame: On Delegation and Responsibility. Review of Economic Studies 79 (1), 67–87.

Battigalli, B. P., Dufwenberg, M., 2007. Guilt in Games. AEA Papers and Proceedings 97 (2).

Becker, S. O., Boeckh, K., Hainz, C., Woessmann, L., 2016. The Empire Is Dead, Long Live the Empire! Long-Run Persistence of Trust and Corruption in the Bureaucracy. The Economic Journal 126 (590), 40 – 74.

Bem, D. J., 1972. Self-perception theory. In: Advances in experimental social psychology. Vol. 6. Elsevier, pp. 1–62.

Bénabou, R., Tirole, J., 2006. Incentives and prosocial behavior. American Economic Review 96, 1652–1678.

Bénabou, R., Tirole, J., 2011. Laws and Norms. Working Paper 17579, 1–44.

Bénabou, R., Tirole, J., 2016. Mindful Economics: The Production, Consumption, and Value of Beliefs. Journal of Economic Perspectives 30 (3), 141–164.

Bennett, M., Yeeles, C., 1990. Children's Understanding of Showing Off. The Journal of Social Psychology 130 (5), 591–596.

Bernheim, B. D., 1994. A Theory of Conformity. Journal of Political Economy 102 (5), 841–877.

Bhanot, S. P., 2017. Rank and response: A field experiment on peer information and water use behavior. Journal of Economic Psychology 62 (May), 155–172.

Bicchieri, C., Xiao, E., 2009. Do the right thing: But only if others do so. Journal of Behavioral Decision Making 22 (2), 191–208.

Blair, C., Razza, R. P., 2007. Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. Child Development 78 (2), 647–663.

Blake, P. R., McAuliffe, K., Warneken, F., 2014. The developmental origins of fairness: The knowledge-behavior gap. Trends in Cognitive Sciences 18 (11), 559–561.

Blanco, M., Engelmann, D., Koch, A. K., 2014. Preferences and Beliefs in a Sequential Social Dilemma: A Within-Subjects Analysis. Games and Economic Behavior 87, 122–135.

Bodner, R., Prelec, D., 2003. Self-signaling in a neo-Calvinist model of everyday decision making. In: Psychology of economic decisions. Oxford University Press, London, UK, pp. 105–126.

Böhm, R., Rockenbach, B., Zimmermann, J., 2018. United we stand, divided we fall: The limitations of between-group comparisons for fostering within-group cooperation. Journal of Economic Psychology 69, 19–29.

Böhm, R., Rockenbach, B., Zimmermann, J., 2019. The driving forces behind information acquisition in social decisions.

Böhm, R., Rothermund, K., Kirchkamp, O., 2013. Social categorization and group-motivated interindividual – intergroup discontinuity. European Review of Social Psychology 49 (October 2012), 40–49.

Bolton, G. E., Ockenfels, A., 2000. ERC: A theory of equity, reciprocity, and competition. American Economic Review 90 (1), 166–193.

Bowles, S., Gintis, H., 2008. Cooperation. In: Blume, L., Durlauf, S. (Eds.), The New Palgrave Dictionary of Economics. Macmillan.

Boyd, R., Richerson, P. J., Henrich, J., 2011. The cultural niche: why social learning is essential for human adaptation. Proceedings of the National Academy of Sciences of the United States of America 108, 10918–10925.

Bradley, A., Lawrence, C., Ferguson, E., 2018. Does observability affect prosociality ? Proceedings of the Royal Society B: Biological Sciences 285 (20180116).

Brewer, M. B., 1991. The Social Self: On Being the Same and Different at the Same Time. Journal of personality and social psychology 17 (5), 475482.

Brewer, M. B., Weber, J. G., 1994. Self-evaluation effects of interpersonal versus intergroup social comparison. Journal of personality and social psychology 66 (2), 268–275.

Brocas, I., Carrillo, J. D., 2018a. Iterative dominance in young children: Experimental evidence in simple two-person games. Journal of Economic Behavior & Organization.

Brocas, I., Carrillo, J. D., may 2018b. The determinants of strategic thinking in preschool children. PLOS ONE 13 (5), e0195456.

Burton-Chellew, M. N., West, S. a., 2012. Pseudocompetition among groups increases human cooperation in a public-goods game. Animal Behaviour 84 (4), 947–952.

Butler, R., 1989. Interest in the Task and Interest in Peers ' Work in Competitive and Noncompetitive Conditions : A Developmental Study. Child development 60 (3), 562–570.

Butler, R., 1998. Age Trends in the Use of Social and Temporal Comparison for Self-Evaluation: Examination of a Novel Developmental Hypothesis. Child Development 69 (4), 1054–1073.

Buttelmann, D., Böhm, R., 2014. The Ontogeny of the Motivation That Underlies In-Group Bias. Psychological Science 25 (4), 921–927.

Cárdenas, J. C., Mantilla, C., 2015. Between-group competition, intra-group cooperation and relative performance. Frontiers in Behavioral Neuroscience 9 (February), 1–9.

Carpenter, J., Matthews, P. H., Schirm, J., 2010. Tournaments and office politics: Evidence from a real effort experiment. American Economic Review 100 (1), 504–17.

Cason, T., Mui, V., 1998. Social Influence in the Sequential Dictator Game. Journal of mathematical psychology 42 (2/3), 248–265.

Charness, G., Rabin, M., 2002. Understanding Social Preferences with Simple Tests*. Quarterly Journal of Economics 117 (3), 817–869.

Chaudhuri, A., Graziano, S., 2006. Social Learning and Norms in a Public Goods Experiment with Inter-Generational Advice. Review of Economic Studies 73 (2), 357–380.

Chen, Y., Harper, F. M., Konstan, J., Li, S. X., 2010. Social comparisons and contributions to online communities: A field experiment on MovieLens. American Economic Review 100 (4), 1358–1398.

Chevallier, C., Molesworth, C., Happé, F., 2012. Diminished social motivation negatively impacts reputation management: Autism spectrum disorders as a case in point. PLoS ONE 7 (1), 1–6.

Cialdini, R. B., De Nicholas, M. E., 1989. Self-presentation by association. Journal of personality and social psychology 57 (4), 626.

Cialdini, R. B., Goldstein, N. J., 2004. Social Influence: Compliance and Conformity. Annual Review of Psychology 55 (1), 591–621.

Cialdini, R. B., Kallgren, C. A., Reno, R. R., 1991. A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior. Advances in Experimental Social Psychology 24 (C), 201–234.

Cialdini, R. B., Richard J., B., Avril, T., Marcus Randall, W., Stephen, F., Lloyd Reynolds, S., 1976. Basking in reflected glory: Three (football) field studies. Journal of Personality and Social Psychology 34 (3), 366–375.

Clark, A. E., Senik, C., 2010. Who compares to Whom? The anatomy of income comparisons in Europe. Economic Journal 120 (544), 573–594.

Collins, R. L., 1996. For better or worse: The impact of upward social comparison on self-evaluations. Psychological Bulletin 119 (1), 51–69.

Conrads, J., Irlenbusch, B., 2013. Strategic ignorance in ultimatum bargaining. Journal of Economic Behavior and Organization 92, 104–115.

Corcoran, K., Crusius, J., Mussweiler, T., 2011. Social Comparison : Motives , Standards , and Mechanisms. Theories in Social Psychology (5), 119–139.

Corcoran, K., Mussweiler, T., 2009. Comparative thinking styles in group and person perception: One mechanism–many effects. Social and Personality Psychology Compass 3, 244–259.

Crocker, J., Canevello, A., 2008. Creating and undermining social support in communal relationships: the role of compassionate and self-image goals. Journal of personality and social psychology 95 (3), 555–575.

Croson, R., Shang, J., 2008. The impact of downward social information on contribution decisions. Experimental Economics 11, 221–233.

Damgaard, M. T. D., Gravert, C., 2013. The hidden costs of nudging: Experimental evidence from reminders in fundraising 2473 (650), 2000–2010.

Dana, J., Cain, D. M., Dawes, R. M., 2006. What you don't know won't hurt me: Costly (but quiet) exit in dictator games. Organizational Behavior and Human Decision Processes 100, 193–201.

Dana, J., Weber, R. a., Kuang, J. X., 2007. Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. Economic Theory 33 (1), 67–80.

Danilov, A., Biemann, T., Kring, T., Sliwka, D., 2013. The dark side of team incentives: Experimental evidence on advice quality from financial service professionals. Journal of Economic Behavior and Organization 93, 266–272.

Dellavigna, S., List, J. A., Malmendier, U., 2012. Testing for altruism and social pressure in charitable giving. Quarterly Journal of Economics 127 (1), 1–56.

Diamond, A., 2013. Executive Functions. Annual Review of Psychology 64 (1), 135–168.

Doosje, B., Ellemers, N., Spears, R., 1995. Perceived Intragroup Variability as a Function of Group Status and Identification. Journal of Experimental Social Psychology 31 (5), 410–436.

Dutton, D., 2009. The art instinct: Beauty, pleasure, & human evolution. Oxford University Press, USA.

Eckel, C. C., Wilson, R. K., 2007. Social learning in coordination games: Does status matter? Experimental Economics 10, 317–329.

Engelmann, J. M., Herrmann, E., Tomasello, M., 2012. Five-Year Olds, but Not Chimpanzees, Attempt to Manage Their Reputations. PLoS ONE 7 (10).

Engelmann, J. M., Herrmann, E., Tomasello, M., 2016. The effects of being watched on resource acquisition in chimpanzees and human children. Animal Cognition 19 (1), 147–151.

Engelmann, J. M., Over, H., Herrmann, E., Tomasello, M., 2013. Young children care more about their reputation with ingroup members and potential reciprocators. Developmental Science 16 (6), 952–958.

Exley, C. L., 2016. Excusing selfishness in charitable giving: The role of risk. Review of Economic Studies 83 (2), 587–628.

Falk, A., Fehr, E., Fischbacher, U., 2008. Testing theories of fairness: Intentions matter. Games and Economic Behavior 62 (1), 287–303.

Falk, A., Szech, N., 2014. Diffusion of Being Pivotal and Immoral Outcomes. DIW Berlin (September), 1–28.

Fehr, E., Bernhard, H., Rockenbach, B., 2008. Egalitarianism in young children. Nature 454 (7208), 1079–1083.

Fehr, E., Fischbacher, U., 2004. Third-party punishment and social norms. Evolution and Human Behavior 25 (2), 63–87.

Fehr, E., Schmidt, K. M., 1999. A theory of fairness, compeition, and cooperation. The Quarterly Journal of Economics 114 (3), 817–867.

Fellner, G., Sausgruber, R., Traxler, C., 2013. Testing Enforcement Strategies in the Field: Appeal, Moral Information, Social Information. Journal of the European Economic Association (June 2011), 41.

Festinger, L., 1954. A Theory of Social Comparison Processes. Human Relations 7 (2), 117–140.

Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Experimental Economics 10 (2), 171–178.

Fischbacher, U., Gächter, S., mar 2010. Social preferences, beliefs, and the dynamics of free riding in public goods experiments. American Economic Review 100 (1), 541–556.

Fischbacher, U., Gächter, S., Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. Economics Letters 71 (3), 397–404.

Fong, C. M., Luttmer, E. F. P., 2011. Do fairness and race matter in generosity? Evidence from a nationally representative charity experiment. Journal of Public Economics 95 (5-6), 372–394.

Fong, C. M., Oberholzer-Gee, F., 2011. Truth in giving: Experimental evidence on the welfare effects of informed giving to the poor. Journal of Public Economics 95 (5-6), 436–444.

Frank, R. H., 1985. The demand for unobservable and other nonpositional goods. American Economic Review 75 (1), 101–116.

Frey, B., 1994. Art: The economic point of view. In: Peacock, A., Rizzo, I. (Eds.), Cultural Economics and Cultural Policy. pp. 3–16.

Frey, B. S., Meier, S., 2004. Social Comparisons and Pro-social Behavior: Testing "Conditional Cooperation" in a Field Experiment. The American Economic Review 94 (5), 1717–1722.

Fu, G., Lee, K., 2007. Social grooming in the kindergarten: The emergence of flattery behavior. Developmental Science 10 (2), 255–265.

Fudenberg, D., Levine, D. K., 2006. A Dual Model of Impulse Control. The American Economic Review 96 (5), 1449–1476.

Gächter, S., Nosenzo, D., Renner, E., Sefton, M., 2012. Who makes a good leader? Cooperativeness, optimism, and leading-by-example. Economic Inquiry 50 (4), 953–967.

Gächter, S., Renner, E., 2018. Leaders as role models and 'belief managers' in social dilemmas. Journal of Economic Behavior & Organization 154, 321–334.

Garcia, S. M., Song, H., Tesser, A., 2010. Tainted recommendations: The social comparison bias. Organizational Behavior and Human Decision Processes 113 (2), 97–101.

Garcia, S. M., Tor, a., Schiff, T. M., 2013. The Psychology of Competition: A Social Comparison Perspective. Perspectives on Psychological Science 8, 634–650.

Garon, N., 2016. A review of hot executive functions in preschoolers. Journal of Self-Regulation and Regulation 2 (2), 57–80.

Gerstadt, C. L., Hong, Y. J., Diamond, A., 1994. The relationship between cognition and action: performance of children 3 1/2-7 years old on a Stroop-like day-night test. Cognition 53 (2), 129–153.

Gilbert, D. T., Giesler, R. B., Morris, K. A., 1995. When comparisons arise. Journal of Personality and Social Psychology 69 (2), 227.

Gneezy, U., Rustichini, A., 2004. Gender and Competition at a Young Age. American Economic Review: Papers & Proceedings 94 (2), 377–381.

Goette, L., Huffman, D., Meier, S., 2006. The impact of group membership on cooperation and norm enforcement. American Economic Review 96 (2), 212–216.

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 (1), 101–115.

Goodman, J. K., Cryder, C. E., Cheema, A., 2013. Data Collection in a Flat World: The Strengths and Weaknesses of Mechanical Turk Samples. Journal of Behavioral Decision Making 26 (3), 213–224.

Greiner, B., 2015. Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE. Journal of the Economic Science Association 1 (1), 114–125.

Greiner, B., Ockenfels, A., Werner, P., 2012. The dynamic interplay of inequality and trust—An experimental study. Journal of Economic Behavior & Organization 81 (2), 355–365.

Grossman, Z., 2014. Strategic Ignorance and the Robustness of Social Preferences. Management Science 60 (11), 2659–2665.

Grossman, Z., van der Weele, J., 2017. Self-Image and Willful Ignorance in Social Decisions. Journal of the European Economic Association 15 (1), 173–217.

Gul, F., Pesendorfer, W., 2001. Temptation and Self-Control. Econometrica 69 (6), 1403–1435.

Gürerk, Ö., Lauer, T., Scheuermann, M., 2018. Leadership with individual rewards and punishments. Journal of Behavioral and Experimental Economics 74, 57–69.

Güth, W., Levati, M. V., Sutter, M., van der Heijden, E., 2007. Leading by example with and without exclusion power in voluntary contribution experiments. Journal of Public Economics 91 (5), 1023–1042.

Haisley, E. C., Weber, R. A., 2010. Self-serving interpretations of ambiguity in other-regarding behavior. Games and Economic Behavior 68 (2), 614–625.

Harbring, C., Irlenbusch, B., 2011. Sabotage in Tournaments : Evidence from a Laboratory Experiment Sabotage in Tournaments : Evidence from a Laboratory Experiment 65 (April 2015), 425–441.

Hauser, D. J., Schwarz, N., 2016. Attentive Turkers: MTurk participants perform better on online attention checks than subject pool participants. Behavior Research Methods (February 2016), 400–407.

Heintz, C., Giardini, F., Max, S., 2015. Facing expectations : Those that we prefer to fulfil and those that we disregard. Judgment and Decision Making 10 (5), 442–455.

Herrmann, B., Thöni, C., Gächter, S., 2008. Antisocial punishment across societies. Science (New York, N.Y.) 319, 1362–1367.

Herrmann, E., Engelmann, J. M., Tomasello, M., 2019. Children engage in competitive altruism. Journal of Experimental Child Psychology 179, 176–189.

Horton, J. J., Rand, D. G., Zeckhauser, R. J., 2011. The online laboratory: Conducting experiments in a real labor market. Experimental Economics 14 (3), 399–425.

Huguet, P., Dumas, F., Monteil, J. M., Genestoux, N., 2001. Social comparison choices in the classroom: Further evidence for students' upward comparison tendency and its beneficial impact on performance. European Journal of Social Psychology 31 (5), 557–578.

Isaac, R. M., Walker, J. A., 1988. Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism. The Quarterly Journal of Economics 103 (1), 179–199.

Jehn, K. A., Northcraft, G. B., Neale, M. A., 1999. Why Differences Make a Difference: A Field Study of Diversity, Conflict, and Performance in Workgroups. Administrative Science Quarterly 44 (4), 741.

Jordan, J., Mcauliffe, K., Rand, D., 2015. The effects of endowment size and strategy method on third party punishment. Experimental Economics 19 (4), 741–763.

Kadish, S. H., Schulhofer, S. J., Barkow, R. E., 2016. Criminal law and its processes: cases and materials. Wolters Kluwer Law & Business.

Kahan, D., Peters, E., Wittlin, M., Slovic, P., Ouellette, L., Braman, D., Mandel, G., 2012. The polarizing impact of science literacy and numeracy on perceived climate change risks. Nature Climate Change 2 (10), 732–735.

Kahneman, D., Knetsch, J. L., Thaler, R. H., 1986. Fairness and the assumptions of economics. Journal of business, S285–S300.

Kelley, H. H., Stahelski, A. J., 1970. Social interaction basis of cooperators' and competitors' beliefs about others. Journal of Personality and Social Psychology 16, 66–91.

Knutsson, M., Martinsson, P., Wollbrant, C., 2013. Do people avoid opportunities to donate?. A natural field experiment on recycling and, charitable giving. Journal of Economic Behavior and Organization 93, 71–77.

Konow, J., 2000. Fair Shares : Accountability and Cognitive Dissonance in Allocation Decisions Fair Shares : Accountability and Cognitive Dissonance in Allocation Decisions. The American Economic Review 90 (4), 1072–1091.

Krupka, E., Weber, R. a., 2009. The focusing and informational effects of norms on pro-social behavior. Journal of Economic Psychology 30 (3), 307–320.

Krupka, E. L., Croson, R. T. A., 2016. The differential impact of social norms cues on charitable contributions. Journal of Economic Behavior and Organization 128, 149–158.

Kuang, X. J., Weber, R. a., Dana, J., 2007. How effective is advice from interested parties? Journal of Economic Behavior & Organization 62 (4), 591–604.

Kunda, Z., 1990. The case for motivated reasoning. Psychological bulletin 108 (3), 480–498.

Larson, T., Capra, C. M., 2009. Exploiting moral wiggle room: Illusory preference for fairness? A comment. Judgment and Decision Making 4 (6), 467–474.

Leary, M. R., 2007. Motivational and emotional aspects of the self. Annual review of psychology 58, 317–344.

Leary, M. R., Baumeister, R., 2000. The nature and function of self-esteem: Sociometer theory. Advances in Experimental Social Psychology 32, 1–62.

Levati, M. V., Sutter, M., Heijden, E. V. D., Levati, M. V., Sutter, M., 2007. Leading by Example in a Public Goods Experiment with Heterogeneity and Incomplete Information for groups in which they have to Individuals often in and. Journal of Conflict Resolution 51 (5), 793–818.

Linardi, S., McConnell, M. A., 2011. No excuses for good behavior: Volunteering and the social environment. Journal of Public Economics 95 (5-6), 445–454.

Luban, D., 1999. Contrived ignorance. Georgetown Law Journal 87 (4), 957.

Magid, R. W., Schulz, L. E., 2010. Quit while you ' re ahead: Preschoolers ' persistence and willingness to accept challenges are affected by social comparison. CogSci.

Marsh, H. W., Kong, C. K., Hau, K.-T., 2000. Longitudinal multilevel models of the big-fish-little-pond effect on academic self-concept: counterbalancing contrast and reflected-glory effects in Hong Kong schools. Journal of personality and social psychology 78 (2), 337–349.

Marsh, H. W., Parker, J. W., 1984. Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? Journal of personality and social psychology 47 (1), 213.

Mazar, N., Amir, O. N., Ariely, D. A. N., 2008. The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. Journal of Marketing Research XLV (December), 633–644.

McGraw, K. M., 1991. Managing Blame : An Experimental Test of the Effects of Political Accounts. American Political Science Review 85 (4), 1133–1157.

Meer, J., 2011. Brother, can you spare a dime? Peer pressure in charitable solicitation. Journal of Public Economics 95 (7-8), 926–941.

Mischel, W., Shoda, Y., Rodriguez, M. I., 1989. Delay of gratification in children. Science 244 (4907), 933–938.

Mitzkewitz, M., Nagel, R., 1993. Experimental results on ultimatum games with incomplete information. International Journal of Game Theory 22 (2), 171–198.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Roberts, B. W., Ross, S., Sears, M. R., Thomson, W. M., Caspi, A., 2011. A gradient of childhood self-control predicts health, wealth, and public safety. Proceedings of the National Academy of Sciences of the United States of America 108 (7), 2693–2698.

Mosatche, H. S., Bragonier, P., 1981. An Observational Study of Social Comparison in Preschoolers. Child Development 52 (1), 376–378.

Mussweiler, T., 2003. Comparison processes in social judgment: mechanisms and consequences. Psychological review 110 (3), 472–489.

Norgaard, K. M., 2006. "People want to protect themselves a little bit": Emotions, denial, and social movement nonparticipation. Sociological Inquiry 76 (3), 372–396.

Ockenfels, A., Werner, P., 2012. 'Hiding behind a small cake' in a newspaper dictator game. Journal of Economic Behavior & Organization 82 (1), 82–85.

Pomerantz, E. M., Ruble, D. N., Frey, K. S., Greulieh, F., 1995. Meeting Goals and Confronting Conflict: Children ' s Changing Perceptions of Social Comparison. Child Development 66 (3), 723–738.

Potters, J., Sefton, M., Vesterlund, L., 2007. Leading-by-example and signaling in voluntary contribution games: an experimental study. Economic Theory 33 (1), 169–182.

Prendergast, C., Topel, R., 1993. Discretion and bias in performance evaluation. European Economic Review 37 (2-3), 355–365.

Pyszczynski, T., Greenberg, J., LaPrelle, J., 1985. Social comparison after success and failure: Biased search for information consistent with a self-serving conclusion. Journal of Experimental Social Psychology 21 (2), 195–211.

Raihani, N. J., Mace, R., Lamba, S., 2013. The Effect of $1, $5 and $10 Stakes in an Online Dictator Game. PLoS ONE 8 (8), 3–8.

Raihani, N. J., McAuliffe, K., 2014. Dictator game giving: The importance of descriptive versus injunctive norms. PLoS ONE 9 (12), 1–17.

Reinstein, D., Riener, G., 2012. Reputation and influence in charitable giving: An experiment. Theory and Decision 72 (2), 221–243.

Rhodes, M., Brickman, D., 2008. Preschoolers' responses to social comparisons involving relative failure. Psychological Science 19 (10), 968–972.

Rivas, M. F., Sutter, M., 2008. The dos and don'ts of leadership in sequential public goods experiments. Tech. rep., Working Papers in Economics and Statistics.

Rivas, M. F., Sutter, M., 2011. The benefits of voluntary leadership in experimental public goods games. Economics Letters 112 (2), 176–178.

Ruble, D. N., Boggiano, A. K., Feldman, N. S., Loebl, J. H., 1980. Developmental analysis of the role of social comparison in self-evaluation. Developmental Psychology 16 (2), 105–115.

Ruble, D. N., Eisenberg, R., Higgins, E. T., 1994. Developmental Changes in Achievement Evaluation: Motivational Implications of Self-Other Differences. Child Development 65 (4), 1095–1110.

Ruble, D. N., Feldman, N. S., Boggiano, A. K., 1976. Social comparison between young children in achievement situations. Developmental Psychology 12 (3), 192–197.

Rudolph, K. D., Caldwell, M. S., Conley, C. S., mar 2005. Need for Approval and Children's Well-Being. Child Development 76 (2), 309–323.

Sanders, M., 2017. Social influences on charitable giving in the workplace. Journal of Behavioral and Experimental Economics 66, 129–136.

Sausgruber, R., 2009. A note on peer effects between teams. Experimental Economics 12 (October 2006), 193–201.

Schotter, A., Sopher, B., 2003. Social Learning and Coordination Conventions in Intergenerational Games: An Experimental Study. Journal of Political Economy 111 (3), 498–529.

Schotter, A., Sopher, B., 2006. Trust and trustworthiness in games: An experimental study of intergenerational advice. Experimental Economics 9 (2), 123–145.

Schram, A., Charness, G., 2015. Inducing Social Norms in Laboratory Allocation Choices. Management Science 61 (7), 1531–1546.

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., Griskevicius, V., 2007. The constructive, destructive, and reconstructive power of social norms: Research article. Psychological Science 18 (5), 429–434.

Shang, J., Croson, R., 2009. A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods. The Economic Journal 119 (Section 1), 1422–1439.

Shaw, A., Olson, K., 2014. Fairness as partiality aversion: The development of procedural justice. Journal of Experimental Child Psychology 119, 40–53.

Sher, I., Koenig, M., Rustichini, A., 2014. Children's strategic theory of mind. Proceedings of the National Academy of Sciences of the United States of America 111 (37), 13307–12.

Spiekermann, K., Weiss, A., 2016. Objective and subjective compliance: A norm-based explanation of 'moral wiggle room'. Games and Economic Behavior 96 (5), 170–183.

Stipek, D. J., Recchia, S., McClintic, S., 1992. Self-Evaluation in Young Children. Monographs of the Society for Research in Child Development 57 (1), 381–387.

Suri, S., Watts, D. J., 2011. Cooperation and Contagion in Web-Based, Networked Public Goods Experiments. PLoS ONE 6 (3), e16836.

Sutter, M., Glätzle-Rützler, D., oct 2015. Gender Differences in the Willingness to Compete Emerge Early in Life and Persist. Management Science 61 (10), 2339–2354.

Sutter, M., Kocher, M. G., Glätzle-Rützler, D., Trautmann, S. T., February 2013. Impatience and uncertainty: Experimental decisions predict adolescents' field behavior. American Economic Review 103 (1), 510–31.

Sutter, M., Yilmaz, L., Oberauer, M., 2015. Delay of gratification and the role of defaults—An experiment with kindergarten children. Economics Letters 137, 21–24.

Sutter, M., Zoller, C., Glätzle-Rützler, D., 2019. Economic behavior of children and adolescents – A first survey of experimental economics results. European Economic Review 111, 98–121.

Tabellini, G., 2008. The Scope of Cooperation: Values and Incentives. Quarterly Journal of Economics 123 (3), 905–950.

Tajfel, H., Billig, M. G., Bundy, R. P., Flament, C., 1971. Social categorization and intergroup behaviour. European Journal of Social Psychology 1, 149–178.

Tajfel, H., Turner, J., 1979. An Integrative Theory of Intergroup Conflict.

Tan, J. H. W., Bolle, F., 2007. Team competition and the public goods game. Economics Letters 96, 133–139.

Taylor, S. E., Wayment, H. A., Carrillo, M., 1996. Social comparison, self-regulation, and motivation.

Tesser, A., 1988. Toward a Self-evaluation Maintenance Model of Social Behavior. Advances in Experimental Social Psychology 21, 181–227.

Tesser, A., Campbell, J., Smith, M., 1984. Friendship choice and performance: Self-evaluation maintenance in children.

Turner, J. C., 1975. Social comparison and social identity: Some prospects for intergroup behavior. European journal of social psychology 5 (1), 5–34.

Verduyn, P., Lee, D. S., Park, J., Shablack, H., Orvell, A., Bayer, J., Ybarra, O., Jonides, J., Kross, E., 2015. Passive Facebook Usage Undermines Affective Well-Being: Experimental and Longitudinal Evidence. Journal of experimental psychology. General 144 (2), 480–488.

Zelazo, P. D., Carlson, S. M., 2012. Hot and Cool Executive Function in Childhood and Adolescence: Development and Plasticity. Child Development Perspectives 6 (4), 354–360.

Zhong, C.-B., Bohns, V. K., Gino, F., 2010. Good lamps are the best police: Darkness increases dishonesty and self-interested behavior. Psychological science 21 (3), 311–314.

Zimmermann, J., 2019. The desire to set an example: Do people cooperate more if they have social influence?

Zimmermann, J., Buttelmann, D., Böhm, R., Rockenbach, B., 2019. Social-Comparison Engineering in Children.

Erklärung nach §8 der Promotionsordnung vom 17. Februar 2015

Hiermit versichere ich an Eides Statt, dass ich die vorgelegte Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Aussagen, Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise entgeltlich/unentgeltlich (zutreffendes bitte unterstreichen) geholfen:

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Dissertation nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Die Dissertation wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt. Ich versichere, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.

Unterschrift:

Curriculum Vitae

Personal Details

Name:	Jarid Emanuel Zimmermann
Date of Birth:	21st April, 1989
Place of Birth:	Nelson, Canada

Education

2015 - 2019	PhD Student, University of Cologne, Cologne
2012 - 2014	Master of Arts, Economics, University of Zurich, Zurich
2009 - 2012	Bachelor of Science, Global Business Management, University of Augsburg, Augsburg
2008	Allgemeine Hochschulreife, Hermann-Billung-Gymnasium, Celle