

Determinants of Norm Compliance: Moral Similarity and Group Identification*

Alexander Schneeberger^{ab}

Erin L. Krupka^{cd}

January 21, 2025

Abstract

What determines whether someone complies with a social norm? The social identity approach offers a mechanism for norm compliance: a person who feels similar to a group identifies more with that group and, in turn, complies with the group's norms. Using an experiment, we test whether similarity in values increases identification and adherence to a group rule. To do so, we varied the similarity/dissimilarity between the values of an individual and members of a social group and measured group identification and rule compliance. We find that similarity in values increased group identification, and group identification increased rule compliance. We show that this behavior change was due to increased group norm sensitivity rather than changes in norms to follow rules when they come from similar or dissimilar groups. We advance the study of social identity by establishing a causal pathway between group identification and behavior change. We also contribute to the management literature by showing that aligning organizational values with those of the workforce is a viable and implementable mechanism for increasing policy and guideline adherence.

Keywords: Moral Similarity, Moral Foundations Theory, Group Identification, Group Norm Sensitivity

JEL Classification Numbers: *C91, D01*

***IRB:** This research was approved by the University of Michigan - Health Sciences and Behavioral Sciences Institutional Review Board (HUM00163304). **Contact:** schneeberger@wiso.uni-koeln.de (A. Schneeberger) and ekrupka@umich.edu (E. L. Krupka).

^aUniversity of Cologne, Universitätsstraße 22a, 50937 Cologne, Germany.

^bMax Planck Institute for Research on Collective Goods, Kurt-Schumacher-Straße 10, 53113 Bonn, Germany.

^cSchool of Information, University of Michigan, 105 S. State St., Ann Arbor, MI 48109, USA.

^dInstitute for the Study of Labor (IZA), Schaumburg-Lippe-Straße 5–9, 53113 Bonn, Germany.

Acknowledgement

We thank the Max-Planck Institute for Research on Collective Goods for financial support. We also thank the participants of seminars at the Max-Planck Institute, the University of Cologne, and the IMPRS Uncertainty for valuable feedback. Lastly, we thank Erik Kimbrough for providing data for our power calculations.

1 Introduction

In today’s workplaces, group work is ubiquitous (Lazear & Shaw, 2007). However, for it to be effective, it often requires group members to follow their group’s implicit and explicit rules even when those conflict with more self-serving incentives. The social identity approach has received considerable attention as a way to model the behavior of individuals in social groups (Hornsey, 2008; Tajfel & Turner, 1979; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987; Turner & Reynolds, 2012). The basic idea is that individuals are members of social groups¹ and gain utility from complying with, or disutility from violating, group-specific social norms (Akerlof & Kranton, 2000; Chang, Chen, & Krupka, 2019; Krupka, Leider, & Jiang, 2017). This framework provides explanations for many organizational behaviors that traditional economic theories can not easily explain (Akerlof & Kranton, 2000, 2002, 2005, 2008).² However, most empirical studies identified the average effect of social identities on behavior, even though many theories assume (Akerlof & Kranton, 2000) and some empirical evidence shows (Kimbrough & Vostroknutov, 2016, 2018) that individuals vary in how strongly a particular social identity might impact their behavior. One source of this individual heterogeneity may be the extent to which an actor identifies with the social group (Hornsey, 2008; Tajfel & Turner, 1979; Turner et al., 1987; Turner & Reynolds, 2012).³

In this study, we experimentally test whether higher group identification influences behavior *through* increased group norm sensitivity. The term “(group) norm sensitivity” is used here to refer to the weight individuals place on their (group-specific) norm-dependent utility.⁴ We manipulated group identification by exogenously varying the similarity in moral preferences between an individual and an unknown social

¹According to Tajfel and Turner (1979), “We can conceptualize a group ... as a collection of individuals who perceive themselves to be members of the same social category, share some emotional involvement in this common definition of themselves, and achieve some degree of social consensus about the evaluation of their group and of their membership of it.”

²Others have also suggested that social identity, and the identity associated norms, might help to better understand social preferences (Kimbrough & Vostroknutov, 2016; Krupka & Weber, 2013), peer effects (Gächter, Gerhards, & Nosenzo, 2017), and framing (Chang et al., 2019).

³Tropp and Wright (2001) defined identification as “the degree to which the ingroup is included in the self.”

⁴In other words, norm sensitivity describes an individual’s preference to comply with the normative expectations of the general population (e.g., the subject pool). Group norm sensitivity describes an individual’s preference to comply with the normative expectations of a specific subset of it (e.g., democrats or republicans in the subject pool). Both are distinct from norm compliance, which describes a behavior that other incentives might have influenced.

group.⁵ To characterize the morality of individuals, we drew from moral foundations theory (Graham et al., 2013; Graham et al., 2011; Haidt & Joseph, 2004) and used the “progressivism” scale (Clark et al., 2017). Progressivism measures the degree to which individuals endorse moral values that focus on the individual over moral values that bind people into larger groups. In one of our treatments, the subjects’ and the social group’s progressivism scores were similar, and in another, they were dissimilar. Subsequently, we measured participants’ group identification and costly compliance with a rule created by the similar/dissimilar social group. We tested whether greater moral similarity between an individual and a social group increases group identification and subsequent compliance with that group’s normative expectations. We find that moral similarity in progressivism increases group identification, and group identification increases compliance with rules from the social group. We show that this behavior change was due to increased group norm sensitivity rather than differences in the group norms.

Our contribution to the economics literature on social identity is to establish a causal mechanism from identification to group norm sensitivity. Prior work in economics examined how group identification affects behavior (Chen & Chen, 2011; Chen & Li, 2009; Goette, Huffman, & Meier, 2006), but that work has not disentangled whether the behavioral effect was due to a change in group norm sensitivity or due to a change in the group norms themselves. We disentangle these mechanisms by manipulating group identification without affecting the relevant group norms. In doing so, we shed light on why there is heterogeneity in norm compliance and when a social identity is likely to impact behavior.

Our contribution to the management literature is on adherence to organizational rules and norms; we not only provide causal evidence of values alignment and rule following, but also suggest a way to create enduring alignment through value sharing. Organizational policies or guidelines that are difficult to enforce but widely adopted - such as “customer focused service” or policies governing integrity and discrimination - are the underpinnings of a strong and loyal work force that results in an effective organization. Nonetheless, a persistent challenge is motivating employees to embrace these policies, especially when they are hard to enforce and when financial incentives may discourage their adoption. We show that connecting and aligning corporate values with those of the workforce, creating work groups that share common values or giving work groups time to develop those common values, is a viable and implementable mechanism for increasing policy and guideline adherence.

⁵According to research in social psychology, morality rather than competence or sociability is most important for a positive evaluation of the own ingroup (Leach, Ellemers, & Barreto, 2007).

2 Related Literature

Our work is closely related to the social-psychological research on self-categorization theory (Turner et al., 1987).⁶ According to self-categorization theory, individuals segment, classify and order their social environment in a process called categorization. In this process, they put themselves and others into different categories. Crucially, which categorization is chosen depends on the comparative fit of the categorization. Categories form in such a way that they maximize intraclass similarities and interclass differences (This is referred to as the meta-contrast principle). Subsequently, the chosen categorization influences self-perception⁷ and behavior. In particular, individuals who define themselves via a social identity engage in depersonalization and self-stereotypization. They no longer see themselves and others in the category as individuals but as exemplars of the group prototype. This perceptual shift leads to behavior that is compliant with the normative expectations of a group. Expressed differently, Turner and Reynolds (2012) write, “...internalization [of the category] is critical to the emergent social norms having an impact on one’s attitudes and behaviour...*and is affected by the degree to which individuals consider themselves psychologically to be members of the particular group* [emphasis added]”.⁸ In our experiments, self-categorization theory predicts that individuals are more likely to self-categorize into similar rather than dissimilar social groups. This categorization creates group identification and causes individuals to comply with behaviors that are seen as prototypical of the social group (i.e., group norms).

There is prior work which demonstrates group identification can influence behavior. Eckel and Grossman (2005) showed that group identification increases cooperation with groups in repeated public goods games. Goette et al. (2006) showed that individuals are more likely to cooperate with ingroup members in simultaneous prisoner’s dilemmas and that third parties are more likely to punish individuals that harmed an ingroup member in a prisoner’s dilemma with third-party punishment. Bernhard, Fehr, and Fischbacher (2006) examined how group identity in small tribes affects behavior in dictator games with third-party punishment. They found that third parties are more likely to punish dictators that harmed an ingroup recipient,

⁶The social-psychological social identity approach (Hornsey, 2008; Turner & Reynolds, 2012) consists of social identity theory (Tajfel & Turner, 1979) and self-categorization theory (Turner et al., 1987; Turner & Reynolds, 2012). While social identity theory mainly describes intergroup processes (e.g., why do individuals display ingroup favoritism), self-categorization theory focuses on intragroup processes (e.g., what are the processes that lead to categorization and why does categorization change behavior).

⁷Self-categorization may occur at various levels of abstraction: Individuals might define themselves on a subordinate level based on interpersonal comparisons (personal identity), on an intermediate level based on intergroup comparisons (social identity), or on a superordinate level as a human being (human identity).

⁸In subsequent work, Turner (1991) and Turner et al. (1987) showed that what an actor considers factual or accurate is a consequence of being in a group with others whom the actor believes agree with their own responses. Turner and Reynolds (2012) summarized these findings as follows: “Because other ingroup members are viewed as similar to oneself, they become a valid source of information and a testing ground for one’s own views on relevant dimensions.”

that third parties are less likely to punish ingroup dictators that harmed an ingroup recipient, and that dictators transfer more to ingroup recipients. Shang, Reed, and Croson (2008) showed that peer effects are stronger if the peer is an ingroup member rather than an outgroup member. Chen and Li (2009) found that ingroup identification increases ingroup favoritism in other-other allocation tasks. Chen and Chen (2011) show that ingroup identification influences equilibrium selection in minimum-effort games. Lastly, Jiang and Li (2019) examined how identity affects behavior in principle agent games. They found that principals make more generous offers to ingroup rather than outgroup agents, that ingroup rather than outgroup agents are less tolerant of principal’s low offers, and that ingroup rather than outgroup agents are more responsive to generous offers.

Recent research accounts for the effect mostly through changes in group norms. For example, Harris, Herrmann, Kontoleon, and Newton (2015) showed that third parties punish ingroup dictators less for displays of ingroup favoritism compared to outgroup dictators. The authors interpreted the lack of punishment for ingroup dictators as evidence that ingroup favoritism in itself is prescribed by group norms. More recently, Dimant (2024) used an incentivized group norm elicitation to empirically show that ingroup favoring allocations in dictator and public good games are more socially appropriate. Lastly, Kimbrough and Vostroknutov (2020) showed that their theoretical model is able to account for ingroup favoritism through changes in the group norms.

Yet, the effect might also be explained by changes in group norm sensitivity. There exists some direct empirical evidence of heterogeneity in norm sensitivity. Kimbrough and Vostroknutov (2016, 2018) introduce an incentivized experimental task, the rule-following task, that allows researchers to approximate an individual’s norm sensitivity. They show that heterogeneity in norm sensitivity exists and that the measure predicts norm compliance in public goods, trust, dictator, and ultimatum games. To our knowledge, there has been only one study that identified a variable that influences norm sensitivity. Kryszowski and Tremewan (2021) find that anonymity reduces the norm sensitivity of dictators in dictator games. It should be noted that the authors use statistical estimation techniques to determine norm sensitivity. In contrast, in our experiment we modify the task of Kimbrough and Vostroknutov to measure group norm sensitivity and are therefore less reliant on distributional assumptions.

As a consequence, it remains unclear to what degree observed behavioral effects are due to changes in norm sensitivity or changes in the norm. We argue that while group identification might influence behavior by changing the norms directly, the explanation is incomplete and or possibly misidentified because it ignores the important psychological mechanism of identification itself (as theorized by Turner et al. (1987)).⁹ Therefore

⁹Pickup, Kimbrough, and de Rooij (2021) examined how group identification with political identities affects the willingness to incur a cost to prevent donations to identity inconsistent lobbying organizations. In their study, they observed a negative correlation between group identification and the willingness to incur a cost. They argue that this is (at least partly) due to an increase in group norm sensitivity. However, their study also shows that group identification changes norm knowledge and the

we examine if group identification can influence behavior in an experimental task where the group norms are held constant. To do so, we create social groups in a laboratory setting based on similar or dissimilar values.

A common approach to studying how social groups impact behavior is to induce social identities in the laboratory,¹⁰ to recruit to the laboratory and group by naturally occurring identities, or to prime naturally occurring identities.¹¹ In economics, one of the most common way to induce social identities is using art-work preferences (henceforth, the Klee-Kandinsky task). Subjects are assigned to one of two groups based on their reported preference regarding pairs of Klee and Kandinsky paintings.¹² The appeal of this approach is that one can create groups based on a preference that are (theoretically) orthogonal to the preferences or topic of study at hand. We extend this work by creating social groups in the laboratory based on naturally occurring moral values (rather than art work preferences) without compromising identification of treatment effects in our study.¹³

To conceptualize morality, we used moral foundations theory (Graham et al., 2013; Graham et al., 2011; Haidt & Joseph, 2004). The theory draws on evolutionary psychology and anthropology to identify five dimensions or “foundations” of morality with which societies construct their moral values. For this reason, moral foundations theory is particularly suitable to measure and describe differences in moral concerns across individuals, social groups, and cultures. The five moral foundations are harm/care, fairness/reciprocity, ingroup/loyalty, authority/respect, and purity/sanctity (in the following, we will only call them harm, fairness, ingroup, authority, and purity).¹⁴ The first two foundations (harm and fairness) are often called the “individualizing” foundations since they describe moral values concerned with protecting the rights and liberties

perception of the (unincentivized) group norm. As a consequence, it remains unclear to what degree the behavioral effect was due to changes in group norm sensitivity or the perception of the group norm.

¹⁰This is also referred to as induced social identities (see Chen & Li, 2009).

¹¹Naturally occurring identities are such things as political affiliation, race, gender, ethnicity. Induced social identities are typically constructed in the lab and might be as simple as “blue group” and “red group” etc. Priming refers to the practice of making a particular identity salient or more prominent to a person for a short period of time. For a comprehensive review of these methods see Chen and Li (2009) and for examples see Chen and Chen (2011), Kranton, Pease, Sanders, and Huettel (2020), and Kranton and Sanders (2017). Several studies have demonstrated that one’s naturally occurring identity or laboratory induced identities influences behavior in controlled laboratory experiments.

¹²In each pair, there is one painting by Paul Klee and one painting by Wassily Kandinsky. Each subject independently chooses which painting they prefer in each pair without being told the artist of each painting. After all subjects make their decisions, subjects are sorted into groups based on how many Klee paintings they prefer.

¹³Unlike naturally occurring identities, such as political etc., the social groups we create for this study do not “exist” in any formal sense prior to subjects’ participation in our experiment. As such, one can think of our approach as more similar to inducing identities. However, these groupings based on moral values do offer a more natural mapping than the Klee-Kandinsky task to one dimension on which people are likely to sort into groups in their work and personal lives.

¹⁴Iyer, Koleva, Graham, Ditto, and Haidt (2012) have argued for an additional liberty/oppression foundation. For a discussion on the inclusion of this and other foundations, see Graham et al. (2013)

of individuals. In contrast, the last three foundations (authority, ingroup, and purity) are often called the “binding” foundations since they describe moral values that bind individuals into larger groups and institutions. In our study, we primarily use the progressivism index. Progressivism describes the degree to which individuals endorse the individualizing moral foundations over the binding moral foundations (Clark et al., 2017).

Moral foundations theory has been used, among others, by Clark et al. (2017) to show that participants’ measured progressivism is positively correlated with cooperative behavior in prisoner’s dilemmas and trust games. However, other studies examined how moral similarity in moral foundations affects preferences and behavior more directly. Dehghani et al. (2016) showed that higher moral similarity, in particular in the domain of purity, leads to less social distancing in social networks and laboratory experiments. Winterich, Zhang, and Mittal (2012) discovered that moral similarity between an individual and a charity increases donations. Other researchers showed that similarity in moral foundations between voters and political parties predicts the intention to vote (Johnson et al., 2014) and voting decision (Enke, 2020; Franks & Scherr, 2015; Milesi, 2017).

3 Theoretic Framework

In the following, we provide a theoretical framework that motivates our experimental design and empirical strategy. The framework is based on models articulated by Akerlof and Kranton (2000), Kimbrough and Vostroknutov (2016), and Krupka and Weber (2013). In our framework, an individual i is introduced to a new social group g . Subsequently, the individual must choose an action $x \in \mathbb{R}^+$ in a non-strategic decision environment. In our experiment, the action describes an allocation in a modified rule-following task (Kimbrough & Vostroknutov, 2016, 2018). Crucially, while making this decision, the individual is aware of the normative expectation η of the social group and can choose to comply with that expectation. We define the individual’s utility function as follows:

$$U_i(x) = x + \phi_i \eta(x) \tag{1}$$

$$\phi_i = \phi(I_i) \tag{2}$$

$$I_i = I(D(m_i, m_g)) \tag{3}$$

We can write the utility function U_i (see Equation (1)) as consisting of two additively separable¹⁵ parts: The standard and the norm-dependent utility component. In our model, the standard utility component

¹⁵Several authors have used similar additively separable utility functions to model social identity, including Akerlof and Kranton (2002, 2005), Benjamin, Choi, and Strickland (2010), Chang et al. (2019), Kimbrough and Vostroknutov (2016), Krupka et al. (2023), and Krupka and Weber (2013)

contains only the valuation that the individual places on their monetary gain. In doing so, we assume that the individual’s payoff equals the selected action x , and the individual gets linear consumption utility from money.¹⁶

The norm-dependent utility component potentially creates (dis)utility if the individual (violates) follows the group norm of the social group. In particular, the group norm function $\eta : \mathbb{R}^+ \rightarrow [-1, 1]$ assigns to each action x a social appropriateness rating. In doing so, assume that the group norm function $\eta(x)$ increases with the selected action x . In equation 1, the group norm function is weighted by the individual’s sensitivity to the norm, $\phi_i \in \mathbb{R}^+$.¹⁷

Crucially, informed by Turner’s self-categorization theory (Turner, 1991), we model norm sensitivity as a function of the identification an individual has with the respective social group. In particular, we assume that group identification $I_i \in \mathbb{R}^+$ is mapped via the group norm sensitivity function $\phi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ into group norm sensitivity ϕ_i (see Equation (2)). In turn, the distance between one’s own moral position, $m_i \in \mathbb{R}$, and the average moral position of the social group $m_g \in \mathbb{R}$, $D(m_i, m_g)$ maps into the group identification function $I : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ to group identification I_i (see equation (3)).¹⁸ Thus, our theory states that identification is a function of perceived value alignment between one’s self and a social group. Identification, in turn, affects the weight an individual places on complying with the group norm. Lastly, the group norm itself is exogenous.

¹⁶We make the same simplifying assumptions that Kimbrough and Vostroknutov (2018) make to model behavior in the rule-following task: First, we assume linear consumption utility. Second, we assume that monetary gains are zero if the action is zero and increase proportionally to the action. In our and Kimbrough and Vostroknutov’s experiments, the monetary gains were not zero if the action was zero and the monetary gains were not increasing proportionally.

¹⁷Sometimes this concept is referred to as the weight someone places on norm compliance or as the concern for norm adherence. We follow Kimbrough and Vostroknutov (2016) and refer to it as “sensitivity”.

¹⁸One explanation for why individuals identify with social groups that are (morally) similar to them might be that it is rational to do so. (Other researchers have used similar arguments in the past. For example, Bernard, Hett, and Mechtel [2016] and Shayo [2009] assumed in their models that individuals choose their group identification to maximize utility.) A rational individual identifies with a social group if identification increases utility. If an individual does not identify with the social group ($I = 0$ and therefore $\phi_i = 0$), the utility function only includes the standard utility component. However, if the individual identifies with the social group ($I > 0$ and therefore $\phi_i > 0$), the utility function includes both the standard and norm-dependent utility components. This change impacts the utility in two ways: First, identification potentially reduces the utility from the standard utility component since it might change the optimal action. Second, identification potentially increases or decreases the utility from the norm-dependent utility component. A rational individual will identify with a social group if there is at least one action in which the utility gain from the norm-dependent utility component is higher than a potential utility loss from the standard utility component. In other words, the standard and the norm-dependent utility component must be somewhat aligned. Now assume that the individual encounters a social group but is not aware of its normative prescriptions. The similarity between an individual and a social group might signal an alignment of the standard and norm-dependent utility components. Hereby, moral similarity might be perceived as a particularly informative signal. However, other characteristics might also inform the individual. For example, the general similarity in gender, ethnicity, and even art preferences suggests a somewhat related past and might also be used to create expectations about a social group’s normative expectations.

In our experiments, we measured the moral position of the individual m_i , the average moral position of the social group m_g , the identification of the individual with the social group I_i , and the selected action of the individual x . We also followed Krupka and Weber (2013) to obtain an empirical proxy of the norm, $\eta(x)$.

The goal of the experimental design was to distinguish whether group identification (I_i) influences behavior (x) via group norm sensitivity (ϕ_i), while eliminating the possibility that behavior is influenced by changes in the norm function ($\eta(x)$). We followed Kimbrough and Vostroknutov (2016, 2018) in using the rule-following task to approximate individual sensitivity to the group norm, ϕ_i . According to the authors, the observed action x of a utility-maximizing individual has a positive monotonic relationship with the group norm sensitivity ϕ_i . However, for this to be true, the norm function η must be a unique norm shared by all members of both treatments. In other words, our treatment manipulation must not have affected the group norms (articulated in Hypothesis 1). To test this pathway from similarity to increased identification and norm compliance, our experimental design kept the moral position of the individual $m_i \in \mathbb{R}$ fixed while changing the average moral position of the reference social group $m_g \in \mathbb{R}$. We predicted that in the treatment where there was greater moral value alignment between the individual and the assigned group, there would be greater identification with that group (articulated in Hypothesis 2). We predicted that in the treatment where there was greater identification with the group, subjects would be more likely to adhere to the group norm (articulated in Hypothesis 3). Lastly, we predicted that an increase in group identification would lead to increased group norm sensitivity (articulated in Hypotheses 4). We now turn to the design.

4 Experimental Design

We pre-screened all participants for their moral type with the pre-screening survey. From these responses we identified the progressivism score for each person. We then recruited from this subject pool into our study arms: The rule elicitation survey, the choice experiments (with two treatments) and the norms elicitation experiments (with two treatments). An overview of the experimental design can be found in Figure 1. In response to referee comments, we also ran two additional variants of the choice experiment to test the impact of attention check questions on our results (see Appendix C); though not our main focus, we will describe and report on these in the results and discussion section. All instructions can be found in the appendix (see Appendix D). We now describe each part of the design in more detail.

4.1 Pre-screening Survey

In the **pre-screening survey**, our goal was to identify the moral type of all participants and, with this, to create a subject pool where we know the progressivism score for each subject; we used this pool to recruit into our study arms. In the pre-screening survey, participants answered the moral foundations questionnaire (Graham et al., 2011). The moral foundations questionnaire measures the harm, fairness, ingroup, authority, and purity foundations described in moral foundations theory (Graham et al., 2013; Graham et al., 2011;

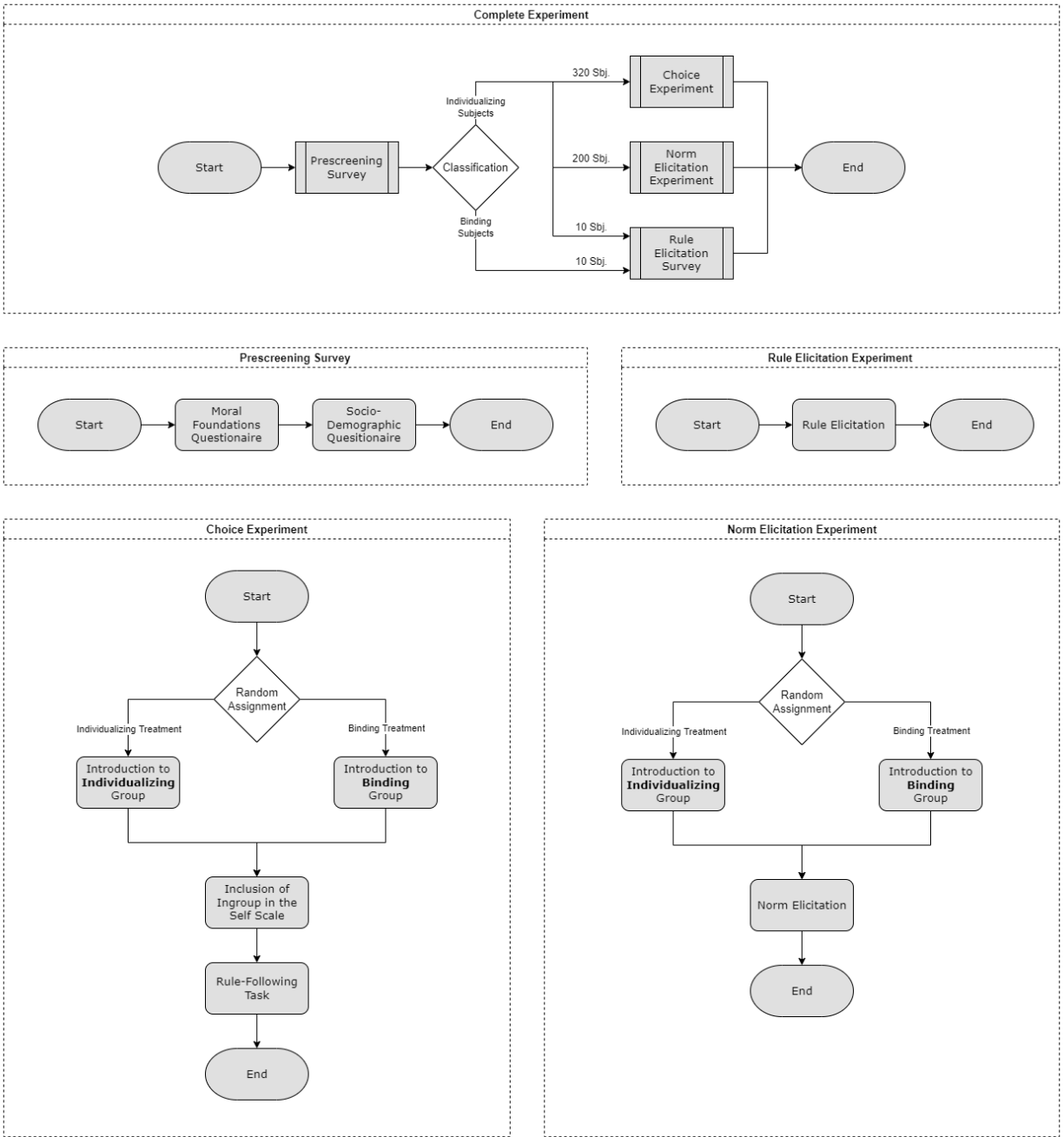


Figure 1: Overview of the Experimental Design

Sbj, Subjects.

Haidt & Joseph, 2004). Each foundation was measured with six items which were scored on a six-point scale ranging from 0 to 5 (see Appendix D.1). Subsequently, we calculated the individualizing and binding indices by averaging the scores of the first two foundations and the last three foundations, respectively. Lastly, we calculated the progressivism index by subtracting the individualizing from the binding index (Clark et al., 2017). This index can take on values ranging from -5 (not at all individualizing & extremely binding) to $+5$ (extremely individualizing & not at all binding). In the following, we will call individuals with a positive progressivism index *individualizing subjects* and individuals with a negative progressivism index *binding subjects*.

4.2 The Rule Elicitation Survey

In the **rule elicitation survey**, our goal was to obtain a statement about the rule in the rule-following task that comes from a social group comprised of either individualizing subjects or binding subjects. Other subjects learned about “the rule” when we used the statement in our choice and norms elicitation experiments.

For the survey, we recruited 10 individualizing subjects and 10 binding subjects from our subject pool. Participants *read* a description of the Kimbrough and Vostroknutov (2018) rule-following task. In the rule-following task, a person must allocate 20 balls into either of two buckets: the yellow bucket or the blue bucket. At the end of the task, the person earns £0.05 for each ball that they place into the blue bucket and £0.10 for each ball that they place into the yellow bucket. While allocating balls into the yellow bucket is payoff dominant, our subjects were told that “*The rule is to put the balls in the blue bucket.*” Our subjects read about this task and were asked to describe “the rule.” They could choose between the following two options: “*The rule is to put the balls in the blue bucket*” and “*The rule is to put the balls in the yellow bucket.*”

This survey may seem odd to administer, but it gives us statements about the rule that we can communicate to other subjects without deception. After collecting the responses, we sorted the participants into individualizing and binding subjects. This allowed us to describe the rule statements to other subjects in our choice and norms elicitation experiments as coming from individualizing or binding subjects. A comparison of the statements between groups found no difference in the modal rule, which always was: “*The rule is to put the balls in the blue bucket*”. From here on, will call the group that consists of only individualizing subjects the “individualizing group” and its rule “the individualizing rule”. Analogously we will call the group that consists of only binding subjects “the binding group” and its rule “the binding rule”. However, we did not use these terms for our subjects; rather the group was simply referred to as “Group A” regardless of whether it was comprised of individualizing or binding subjects.

4.3 The Choice Experiment

The **choice experiment**, and its two treatments, allowed us to measure our dependent variable of interest: following a rule statement to place balls in a bucket when that rule comes from groups composed of similar

or dissimilar others. To do so, we introduced participants to one of two social groups (“the individualizing group” or “the binding group”), determined their level of identification with that group, told them the group’s “rule”, and measured the number of balls they put in the rule-compliant bucket. For the choice experiment we *only recruited individualizing subjects* from our subject pool.¹⁹ The experiment consists of three parts.

Table 1: What Subjects Read to Learn about the Moral Position of Group A

Foundation	Question	Individualizing treatment	Binding treatment
Harm	Whether or not someone cared for someone weak or vulnerable?	Very relevant	Somewhat relevant
Fairness	Whether or not some people were treated differently from others?	Very relevant	Somewhat relevant
Ingroup	Whether or not someone’s action showed love for his or her country?	Not very relevant	Somewhat relevant
Authority	Whether or not someone showed a lack of respect for authority?	Slightly relevant	Somewhat relevant
Purity	Whether or not someone violated standards of purity and decency?	Slightly relevant	Somewhat relevant

Note. Participants only saw the five questions and the corresponding rounded average responses (column 3 or 4 depending on whether they were randomized to the individualizing or binding treatment).

In **part one**, we randomly assigned participants to one of two treatments: The individualizing or the binding treatment. Depending on the treatment, subjects either learned about members who lean individualizing or binding and participated in the rule elicitation survey (in both treatments, simply referred to as “Group A”). To learn about Group A’s members, subjects were shown five questions from the moral foundations questionnaire and the corresponding rounded average response to those questions by members of Group A. Each of the five questions represented one of the five dimensions of moral foundations theory.²⁰

Table 1 shows the rounded average responses of subjects of the individualizing group (column 3) and the binding group (column 4) from our rule elicitation survey. Subjects in our choice experiments were asked to memorize the responses as they subsequently were asked to recall the rounded average response to one randomly selected question. After reviewing the responses, they were shown one randomly chosen (control)

¹⁹We only needed one moral type to test our hypotheses. As such, we chose individualizing subjects.

²⁰We selected the five questions in the following way: For each of the five moral foundations, we chose one item out of the relevance section of the questionnaire. We always chose the item that had the highest factor loading on the particular moral foundation according to the confirmatory factor analysis of Graham et al. (2011).

question and asked to recall the rounded average response for that question from members of “Group A”. If they selected the correct answer, they received £0.50 and otherwise nothing. Participants were informed about whether their recollection was accurate (and they received payment) at the very end when we reported their total earnings.

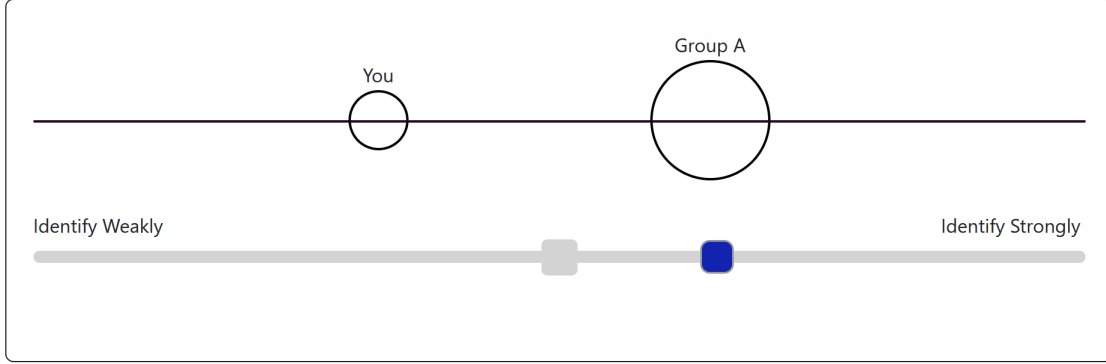


Figure 2: Screenshot of the “*Inclusion of Ingroup in the Self Scale*”

In **part two**, we used a continuous version of the “*Inclusion of Ingroup in the Self Scale*”²¹ (Tropp & Wright, 2001) to measure the degree to which subjects identify with Group A.²² Subjects saw a small circle labeled “You” and a big circle labeled “Group A” (see Figure 2). Participants could move a slider below the circles to the left (moving the circles further apart) or to the right (moving the circles closer to each other).²³ Consistent with prior work, we interpret distant circles as weak, and adjacent circles as strong, identification with Group A. The distance identification measure can take values from 0 (weak identification / the circles are at opposites ends of the scale) to 1 (strong identification / the circles’ centers are at the scale’s midpoint).

In **part three**, participants played a modified version of the rule-following task (Kimbrough & Vostroknutov, 2018), which we call the group rule-following task (see Figure 3, the location of the blue and yellow bucket was left right randomized). While the rule from Kimbrough and Vostroknutov had no obvious source except the experimenter, our rule originates from members of Group A.²⁴ In our study, the rule read: “*According*

²¹The “Inclusion of Ingroup in the Self Scale” is a version of the “Inclusion of Other in Self Scale” (Aron, Aron, & Smollan, 1992). However, instead of indicating the perceived distance to another person, one sets the perceived distance to a social group. The “Inclusion of Other in Self scale” was validated by Gächter, Starmer, and Tufano (2015) and has been used in economics by Bicchieri, Dimant, Gächter, and Nosenzo (2022) and Gächter, Starmer, and Tufano (2017). We use a continuous version of the scale to allow for higher precision. The change is identical to the modification of the discrete to the continuous “Inclusion of Other in Self Scale” (Aron et al., 1992; Le, Moss, & Mashek, 2007).

²²We only measured the identification with the ingroup (Tropp & Wright, 2001) and not the identification with the ingroup and the outgroup (Schubert & Otten, 2002) since we only had a clearly defined ingroup: Group A.

²³We asked participants to “... move the slider to a position in which the distance between the two circles most accurately describes your identification with Group A”.

²⁴In Kimbrough and Vostroknutov the rule read: “The rule is to put the balls in the blue bucket.” Related research demonstrated

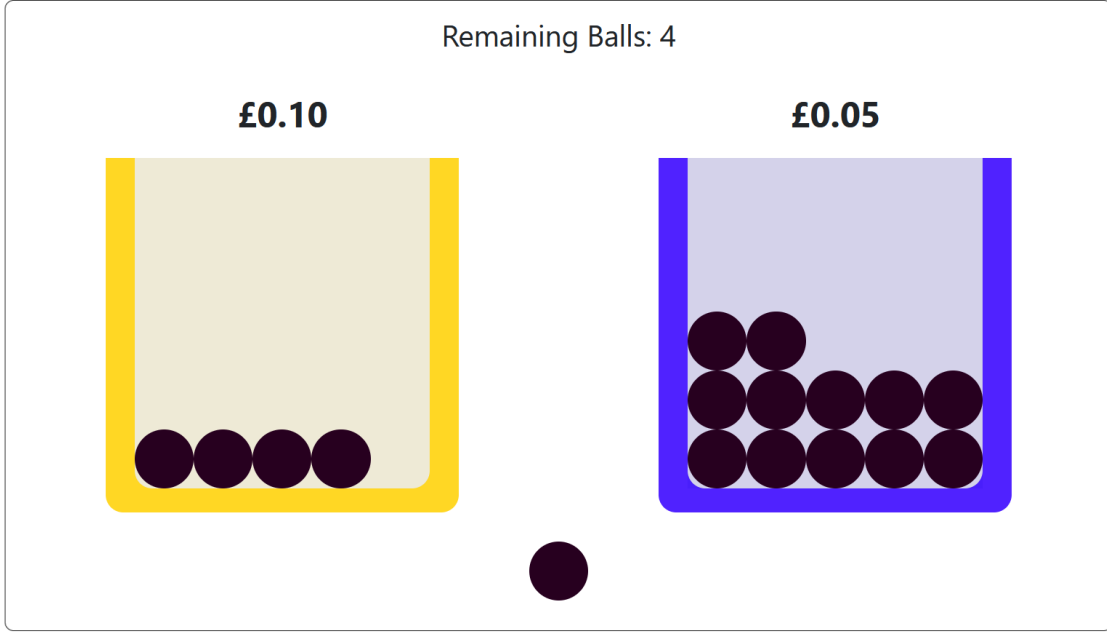


Figure 3: Screenshot of the Group Rule-Following Task. The position of the blue and the yellow bucket were left/right randomized for each subject.

to the members of Group A, the rule is to put the balls in the blue bucket.” The number of balls in the rule-compliant bucket is our dependent variable of interest and our behavioral measure. The experiment concluded with a questionnaire that included an alternative group identification measure used by Chen and Li (2009).

4.4 The Norm Elicitation Experiment

In the **norm elicitation experiment**, our main goal was to obtain data to test whether there were different norms to follow a rule that was given by a similar or dissimilar group (as our identification depends on these norms *not differing*). As in the choice experiment, we randomized these subjects to either be introduced to “Group A” comprised of either individualizing or binding members. As with the choice experiments, we only recruited individualizing subjects from our subject pool. The experiment consisted of two parts.

Part one followed the same process as part one in the choice experiment. In short, we randomly assigned participants to learn of the moral position of the individualizing or the binding “Group A”.

In **part two**, we used the coordination task designed by Krupka and Weber (2013) to elicit the group norms of

that individuals respond to such rule statements. Kimbrough and Vostroknutov (2016, 2018) show that individuals respond to the arbitrary rule of the experimenter, while Pryor, Perfors, and Howe (2019) show that individuals comply with arbitrary descriptive norms even if they do not reflect the preferences of the group.

individualizing subjects for the group rule-following task. We informed participants that they had to evaluate the potential behavior of participants in another study. We described the choice experiment: That other participants were introduced to Group A (the very same group they have just learned about),²⁵ and that the other subjects then had to play the group rule-following task after learning about the rule statement from Group A. Thus, our treatment consisted of whether subjects read about a situation where others learned about the “rule” in the group rule-following tasks from similar or dissimilar others.

Then, our norm elicitation participants played a coordination game with all other participants in the same treatment over appropriateness ratings for each action (“put 0”, “5”, “10”, “15”, and “20” balls into the blue bucket while placing the remaining balls into the yellow bucket). The appropriateness ratings ranged from, “Very socially appropriate,” “Socially appropriate,” “Somewhat socially appropriate,” “Somewhat socially inappropriate,” “Socially inappropriate,” and “Very socially inappropriate.” Recall, that we only recruit individualizing participants to the norm elicitation experiment. To establish common knowledge that they will be coordinating with other individualizing subjects (without explicitly stating this), we tell them how the other participants responded to the moral foundations questions in the pre-screening survey. Specifically, we say that the questions belonged to one of two categories (“category 1” and “category 2”).²⁶ We designed it so that category 1 contained the questions of the individualizing index and category 2 contained the questions of the binding index. Participants then learn that they will try to coordinate ratings in the coordination game with others who stated that questions in category 1 (individualizing) were more important to them than those in category 2.

After all participants completed their ratings, we randomly chose one action (corresponding to an allocation of blue and yellow balls such as “0 in the blue bucket and 20 in the yellow bucket”) and determined the modal social appropriateness rating for that allocation. If a subject selected the same social appropriateness rating as the modal social appropriateness rating of other raters in their treatment, then they received an additional payment of £2. Otherwise, they received nothing.

Krupka and Weber argue that in these coordination games the social norm creates a salient focal point. Consequently, in our experiment individuals have a monetary incentive to reveal their perception of the individualizing group norm about compliance with the rule when it comes from group members who are similar or dissimilar. This allows us to test whether norms to follow the rule differ depending on the composition of the group stating the rule.

Table 2: Number of Subjects and Average Bonus per Experiment

Experiment	Individualizing treatment (n)	Binding treatment (n)	Bonus (£)
Choice experiment	160	160	2.01
Norm elicitation experiment	100	100	0.96

5 Experimental Procedure and Payoffs

We conducted all our surveys and experiments online using the experimental software oTree (Chen, Schonger, & Wickens, 2016) and the online recruitment platform Prolific.com (Palan & Schitter, 2018).²⁷ Hereby, we only invited subjects to the prescreening survey who fulfilled the following requirements: They were United States nationals who were at that time located in the United States, used a desktop or laptop, had between 10 and 1000 previous submissions, and had an approval rate of at least 90%. Furthermore, we only re-invited individualizing participants to the choice and norm elicitation experiment. In doing so, each subject either participated in the rule elicitation survey, the choice experiment, or the norm elicitation experiment. Within the choice and norm elicitation experiment, participants were randomly assigned to either the individualizing or binding treatment.²⁸ For their participation in the prescreening survey, subjects received a participation fee of £1.20. For their participation in the choice or norm elicitation experiment, they received £1.60. In addition, participants earned on average £2.01 in the choice experiment and £0.96 in the norm elicitation experiment (see Table 2). We collected data until we obtained 160 observations in each treatment of the choice experiment and 100 observations in each treatment of the norm elicitation experiment.²⁹

²⁵In other words, if in Part 1 of the norm elicitation experiment, subjects learned about Group A members who had binding values, then they also read about participants in the choice experiment who learned about the same binding Group A members.

²⁶While doing this, we displayed all questions included in either category graphically separated.

²⁷Peer, Brandimarte, Samat, and Acquisti (2017) examined various crowdsourcing platforms and discovered that the data quality on Prolific.com was comparable with Amazon Mechanical Turk. They also found that participants were more naive, less dishonest, and more diverse. Similarly, Douglas, Ewell, and Brauer (2023) found that participants on Prolific.com were more likely to pass attention checks, provide meaningful answers, follow instructions, remember information, have unique IP addresses and geolocation, and spend enough time on pages to read all items than on various other platforms like Amazon Mechanical Turk.

²⁸We did not observe any significant differences in the measured socio-demographic characteristics between the treatments of the choice or norm elicitation experiment (see Appendix B.1).

²⁹The sample size in the choice experiment was based on the following power calculation: We used data from Kimbrough and Vostroknutov (2018) to calculate an expected effect size of 0.3. Subsequently, we used G-Power (Faul, Erdfelder, Lang, & Buchner, 2007) to calculate the minimum required sample size for a one-tailed Wilcoxon-Mann-Whitney test with a parent distribution min ARE, an effect size of 0.3, an α -error probability of 0.05, a power of 0.80, and an even allocation between treatments. Based on this power calculation, we needed at least 160 subjects per treatment. The sample size of the norm

6 Results

We begin by examining the group norms from the norm elicitation experiment and the choices from the choice experiment separately. We conclude by testing the effect of group identification on group norm sensitivity. The analysis was preregistered at The American Economic Association’s registry for randomized controlled trials (<https://doi.org/g5qw>). In Appendix A, we note any deviations from the preregistration.

6.1 Results Norm Elicitation: The Norm to Follow the Rule Did Not Differ by Treatment

We followed Krupka and Weber (2013) in converting the categorical social appropriateness ratings obtained from the norms elicitation experiment into numerical ratings. We coded the rating “Very socially inappropriate” as -1 , the rating “Socially inappropriate” as -0.6 , the rating “Somewhat socially inappropriate” as -0.2 , the rating “Somewhat socially appropriate” as 0.2 , the rating “Socially appropriate” as 0.6 , and finally the rating “Very socially appropriate” as 1 .³⁰ We tested whether receiving the individualizing or binding rule statement in the group rule-following task changed the social appropriateness of allocations. We predicted no significant difference in the group norms, which should prescribe rule following and putting all the balls in the rule-compliant (blue) bucket irrespective of treatment.

Hypothesis 1 (Comparison of the Group Norms): The group norms in the group rule-following task will not differ significantly between the individualizing and the binding treatment.

In Table 3, we report for each action of the group rule-following task the average social appropriateness rating and the frequency distribution of social appropriateness ratings separated by treatment. To test our hypothesis, we used the same approach as Chang et al. (2019). For each of the five actions that we considered in the norm elicitation experiment (place 0, 5, 10, 15, or 20 balls in the rule-compliant bucket), we used a two-sided Wilcoxon-Mann-Whitney test and a two-sided t-test to compare the norm ratings between the individualizing and the binding treatment. Due to the inflated likelihood of type 1 errors, we applied the Bonferroni correction to our statistical significance threshold of $p = 0.05$. The new threshold became $p = 0.01$. As in Chang et al., we considered two group norms to be significantly different from one another when the majority of their norm ratings (in our case three) were significantly different from each other. The p-values of the five t-tests and the five Wilcoxon-Mann-Whitney tests are displayed in Table 3. None of the

elicitation experiment was based on samples used in related studies (Chang et al., 2019).

³⁰The same transformation was used among others by Chang et al. (2019), D’Adda, Drouvelis, and Nosenzo (2016), Erkut, Nosenzo, and Sefton (2015), Gächter, Gerhards, and Nosenzo (2017), Gächter, Nosenzo, and Sefton (2013), Kimbrough and Vostroknutov (2016), and Veselý (2015).

Table 3: Elicited Social Appropriateness Ratings for the Group Rule-Following Task in the Individualizing and Binding Treatment

Action	Individualizing treatment (n = 100)							Binding treatment (n = 100)							P-values ^{ab}		
In blue	Mean ^a	---	---	---	---	+	++	++	++	++	+	+	++	++	++	WMW	T
0 balls	-0.35	44	18	8	8	6	16	-0.22	32	20	8	15	10	15	15	0.134	0.200
5 balls	-0.18	7	42	18	13	13	7	-0.10	9	30	19	16	20	6	6	0.249	0.279
10 balls	0.23	5	5	23	30	19	18	0.21	3	9	23	29	20	16	16	0.741	0.790
15 balls	0.26	5	9	12	23	41	10	0.20	0	13	19	30	32	6	6	0.143	0.329
20 balls	0.36	15	11	6	7	10	51	0.35	12	11	6	12	16	43	43	0.648	0.970

Note. — — —, “very socially inappropriate”; — —, “socially inappropriate”; —, “somewhat socially inappropriate”; +, “somewhat socially appropriate”; ++, “socially appropriate”; ++++, “very socially appropriate”; WMW, Wilcoxon-Mann-Whitney test; T, t-test.

^a We converted the categorical ratings in the same order as described above into the following numerical scores: -1, -0.6, -0.2, 0.2, 0.6, 1.

^b The p-values are based on a two-sided Wilcoxon-Mann-Whitney test and t-test that compared the ratings between treatments.

ten comparisons reached the significance threshold (in all cases $p > 0.134$).³¹ Due to this result, we can approximate group norm sensitivity with rule compliance in the group rule-following task (see 3).

Result 1 (Comparison of the Group Norms): The group norms in the group rule-following task did not differ significantly between the individualizing and the binding treatment.

6.2 Results Choice Experiment: Greater Group Identification in the Individualizing Treatment

We now test whether identification differed by treatment in our choice experiment. In order to test whether our treatment manipulation successfully affected group identification, we compared subjects' group identification in the individualizing and the binding treatment. In line with the psychological social identity approach (Hornsey, 2008; Turner & Reynolds, 2012), we hypothesize that moral similarity increases group identification.

Hypothesis 2 (Comparison of Group Identification): Group identification will be significantly higher in the individualizing treatment than the binding treatment.

We compared responses to the "Inclusion of Ingroup in the Self Scale" between the individualizing and binding treatment. Using responses from how subjects positioned the slider, the self-reported average group identification is 0.80 in the individualizing treatment and 0.62 in the binding treatment. The 0.19 point difference was significant using a one-sided t-test ($p < 0.001$).³²

Result 2 (Comparison of Group Identification): Group identification was significantly higher in the individualizing treatment than in the binding treatment.

6.3 Results Choice Experiment: Higher Group Norm Sensitivity in the Individualizing Treatment

We now test whether our treatment manipulation also affected our participants' group norm sensitivity. Hereby, we use Result 1, and approximate group norm sensitivity with the number of balls in the rule-compliant bucket in the group rule-following task. In line with the psychological social identity approach (Hornsey, 2008; Turner & Reynolds, 2012), we hypothesized that moral similarity would increase group norm sensitivity and, with this, the number of balls placed in the rule-compliant bucket.

³¹In Appendix B.2 we also show that the social appropriateness of an action increases with the number of balls placed in the rule-compliant (blue) bucket.

³²In Appendix B.3, we also considered other identification measures. We did not observe any noticeable differences due to changes in how we measure group identification.

Hypothesis 3 (Comparison of the Number of Balls in the Rule-Compliant Bucket): The number of balls in the rule-compliant bucket will be significantly higher in the individualizing treatment than the binding treatment.

If we exclude those observations assigned to the authority control question ($n=254$), the average participant placed 8.24 and 6.44 balls in the rule-compliant bucket of the individualizing and the binding treatment, respectively. This difference of 1.79 balls was significant according to a one-sided t-test ($p = 0.034$). However, if we use all observations ($n=320$), the participants placed 8.15 and 7.44 balls in the rule-compliant bucket of the individualizing and the binding treatment, respectively. This difference was not significant according to a one-sided t-test ($p = 0.213$).

Though unanticipated in our design, the authority control question (which is one of 5 randomly selected control questions) unexpectedly primed our participants to act more rule-compliant in the binding treatment (see Appendix B.4). This priming effect ran counter to the treatment effect described in Hypothesis 3. Consequently, including observations assigned to the authority control question potentially obscured a treatment effect.

To test Hypothesis 3 without the authority control question, we conducted an additional “limited control questions study” (see Appendix C) where we did not include the authority control question. We recruited 466 new subjects into the pre-screening survey and then randomize them into one of the two treatments in the choice experiment (233 subjects randomized to each treatment arm).³³ Though a description and analysis of the study can be found in Appendix C, we report on rule compliance here. In the “limited control questions study”, participants placed on average 9.41 and 7.79 balls in the rule-compliant bucket of the individualizing and the binding treatment, respectively. The 1.62 balls difference was significant according to both a one-sided t-test ($p = 0.012$) and a Wilcoxon-Mann-Whitney test ($p = 0.009$).

Taken together, the evidence is consistent with Hypothesis 3, that when subject identification is increased with a group, they comply more with the rules of that group. Since Hypothesis 1 demonstrates that the group norms did not change between treatments, we can attribute the increased rule compliance to increased group norm sensitivity. However, the evidence also suggests that the effect can be fragile.

Result 3 (Comparison of the Number of Balls in the Rule-Compliant Bucket): The number of balls in the rule-compliant bucket was significantly higher in the individualizing treatment than the binding treatment if one excludes those observations assigned to the authority control question. If one does not exclude those observations, the effects were directionally as predicted but not significant. Participants in the “limited control questions study” placed significantly more balls in the rule-compliant bucket in the individualizing

³³The study was pre-registered and sample size was based on power calculations which are all reported in Appendix C. We thank referees for the suggestion to collect additional data.

treatment.

6.4 Results: Group Identification Increased Group Norm Sensitivity

In the previous two sections, we demonstrate that both group identification and group norm sensitivity were higher in the individualizing treatment. In this section, we directly examine the relationship between group identification and group norm sensitivity. Again, we use Result 1, and approximate group norm sensitivity with the number of balls in the rule-compliant bucket in the group rule-following task. The psychological social identity approach (Hornsey, 2008; Turner & Reynolds, 2012) suggests that group identification increases the likelihood of prototypical group behavior. As such, we predicted that increased group identification would increase group norm sensitivity and, with this, the number of balls placed in the rule-compliant bucket.

Hypothesis 4 (The Effect of Group Identification on Norm Sensitivity): Higher group identification will increase the number of balls placed in the rule-compliant bucket.

Table 4: The Effect of Group Identification on Rule Compliance

	<i>Balls in the blue bucket</i>			
	(1)		(2)	
	Coef.	Std. err.	Coef.	Std. err.
<i>Identification (distance)</i>	14.131***	(4.034)	14.657***	(4.326)
<i>Binding treatment</i>			-5.036	(4.705)
Constant	-4.636	(3.099)	-3.622	(5.942)
Controls	No		Yes	
Log-likelihood	-738		-725	
BIC	1,493		1,549	
Number of cases	320		320	

Notes. We used Tobit regressions. The controls include the following variables: The subject’s progressivism, the distance in progressivism, one dummy for each control question (except for the harm control question), interactions between the control question dummies and the binding treatment dummy, and the time spent on stages one, two and three. Coef., Coefficient; Std. err., Standard error.

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

To test this, we estimated two Tobit regression models.³⁴ In all models, we used the number of balls placed in the blue (rule-compliant) bucket in the rule-following task as our dependent variable. In the first model, we

³⁴We chose Tobit regression models to handle the censoring at 0 and 20. However, the coefficients of “Identification (distance)” would have been significant and positive in OLS regression models as well (available on request).

regressed the dependent variable against the distance measure of the “Inclusion of Ingroup in the Self Scale.” In the second model, we added a binding treatment dummy and a set of control variables.³⁵ We predicted that group identification would be significant and positive in all specifications. The estimates of the two Tobit regression models can be found in Table 4. As can be seen, the coefficient on group identification is positive and significant in both models ($p < 0.001$). Furthermore, the effect is large in magnitude. If group identification increases from 0 to 1 (on a scale from 0 to 1), the number of balls in the rule-compliant bucket, *ceteris paribus*, increases by approximately 14 balls.³⁶

Result 4 (The Effect of Group Identification on Group Norm Sensitivity): Higher group identification increased the number of balls placed in the rule-compliant bucket.

6.5 Results: Moral Similarity Increased Group Identification

Lastly, we examined how moral similarity to a social group affected group identification. In doing so, we exclusively conceptualized morality by progressivism and moral similarity as the absolute difference between the individual’s progressivism and the social group’s implied progressivism.

On average, participants had a progressivism score of 1.72 in the individualizing treatment and 1.59 in the binding treatment (on a scale from -5 to 5).³⁷ The five answers of the moral foundations questionnaire used to introduce Group A implied a progressivism score of 2.33 in the individualizing treatment and 0 in the binding treatment (on a scale from -5 to 5). Finally, the average absolute difference between the individual’s progressivism and the social group’s implied progressivism was 0.97 in the individualizing treatment and 1.59 in the binding treatment (on a scale from 0 to 10).

To examine how moral similarity affected group identification, we estimated two Tobit regression models with robust standard errors.³⁸ In all models, we used the self-reported distance group identification measure of the “Inclusion of Ingroup in the Self Scale” as our dependent variable. In the first model, we regressed the dependent variable against the subject’s progressivism and the absolute difference between the individual’s progressivism and the social group’s implied progressivism. In the second model, we added a binding

³⁵The set of control variables are: The subject’s progressivism, the distance in progressivism, one dummy for each control question (except for the harm control question), interactions between the control question dummies and the binding treatment dummy, and the time spent on stages one, two, and three.

³⁶In Appendix B.5, we also consider other identification measures. We do not find any noticeable difference due to changes in the conceptualization of group identification.

³⁷By design, our randomization to treatment should and does predict that the difference in the participant’s progressivism was not significant in a one sided t-test (see Appendix B.1).

³⁸We use Tobit regression models to handle the censoring at 0 and 1. However, the coefficients of “Distance progressivism” is also significant and negative in OLS regression models as well (available on request).

treatment dummy and a set of control variables.³⁹

Table 5: The Effect of Moral Similarity on Group Identification

	<i>Identification (distance)</i>			
	(1)		(2)	
	Coef.	Std. err.	Coef.	Std. err.
<i>Progressivism</i>	0.019	(0.015)	0.006	(0.016)
<i>Distance progressivism</i>	-0.118***	(0.017)	-0.081***	(0.020)
<i>Binding treatment</i>			-0.196**	(0.067)
Constant	0.836***	(0.034)	0.890***	(0.062)
Controls	No		Yes	
Log-likelihood	-41		-25	
BIC	104		136	
Number of cases	320		320	

Notes. We used Tobit regressions with robust standard errors. The controls include the following variables: One dummy for each control question (except for the harm control question), interactions between the control question dummies and the binding treatment dummy, and the time spent on stages one and two. Coef., Coefficient; Std. err., Standard error.

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

The results can be found in Table 5. As can be seen, distance progressivism was significant and negative in both specifications ($p < 0.001$). This result suggests that an increase in distance progressivism decreased group identification. In particular, if distance progressivism increased by the average amount observed in our choice experiment (1.28), group identification decreased, ceteris paribus, by 0.10 to 0.15 (on a scale from 0 to 1).⁴⁰

³⁹The set of control variables were: One dummy for each control question (except for the harm control question), interactions between the control question dummies and the binding treatment dummy, and the time spent on stages one and two.

⁴⁰In Appendix B.6, we examined which moral foundations drive the effect of moral similarity on group identification. In Appendix B.7, we examined how robust our results are to changes in the conceptualization of moral similarity. We found that the distance in the individualizing and the binding index both influenced group identification. Furthermore, this effect was mainly driven by the distance in the fairness foundation. We also found that the effect of moral similarity on group identification was robust to changes in the conceptualization of moral similarity.

7 Discussion

In our study, we show that an increase in group identification leads to heightened levels of group norm sensitivity. We also show that an increase in moral similarity based on moral foundations theory leads to elevated levels of group identification.

Our findings connect with a nascent literature regarding the stability of norm sensitivity (Kimbrough et al., 2024; Kimbrough & Vostroknutov, 2016, 2018). While (group) norm sensitivity might be stable in the short term, the moral position of individuals (and consequently of the social groups) changes in the medium term (Van de Vyver, Houston, Abrams, & Vasiljevic, 2016). Hence, group identification and (group) norm sensitivity may have limited temporal stability. With this, we experimentally identify a source of variation in (group) norm sensitivity. However, we also have more direct evidence that the effect can be fragile. To this point, we collected additional data in the “no control questions study” which we report on in Appendix C.3. In the “no control questions study”, we removed the control questions which asked subjects to recall the rounded average response of Group A members to one randomly selected moral foundations question (the incentivized control question in the main study was used to direct attention to the moral position of Group A). This treatment allowed us to indirectly and imperfectly test whether specific awareness about moral similarity or dissimilarity with the rule-setting group impacts rule following. We collected 466 new observations (233 randomized into each treatment of the choice experiment) and find that participants placed, on average, 7.74 and 7.08 balls in the rule-compliant bucket of the individualizing and the binding treatment. Though directionally consistent with our main hypothesis, the 0.66 balls difference was insignificant according to a one-sided t-test ($p = 0.172$). The results from our “no control questions study” and the unanticipated priming effect we observed in our main study, suggest that there is still much work to be done to understand when and for how long, sensitivity to group norms is in effect.

Our findings also help understand how group identification affects ingroup favoritism. Kimbrough and Vostroknutov (2020) recently provided a theoretical explanation of why Chen and Li (2009) might have observed ingroup favoritism in their other-other allocation tasks. They argued that two group norms are present if the allocation is between two participants that belong to distinct social identities. Each group norm favors allocating more money to the ingroup member rather than the outgroup member. According to Kimbrough and Vostroknutov, the allocator exclusively considers the group norm of their own social identity while making the allocation decision. We suggest that this explanation might be incomplete since it assumes an extreme case where individuals only comply with the ingroup norm and ignore the outgroup norm. Our results suggest that the degree to which individuals comply with group norms in itself is a function of group identification. Consequently, an increase in group identification might influence behavior through two channels in other-other allocations tasks: Through an increased group norm sensitivity towards the ingroup norm and a change in the applicable ingroup norm that favors allocations to the ingroup member.

Our findings also shed light on why some individuals are generally prone to ingroup favoritism. Kranton

et al. (2020) and Kranton and Sanders (2017) observed that some individuals are “groupy” and engage in ingroup favoritism independent of the considered social group (i.e., minimal groups and political groups). What makes these groupy individuals different? It might be that groupy individuals have high default group norm sensitivity, and the considered social groups have group norms that prescribe ingroup favoritism. Alternatively, groupy individuals might have typical levels of group norm sensitivity, but their perception of group norms might be biased towards ingroup favoritism. Our results suggest that situational factors like the moral similarity to a social group strongly influence group norm sensitivity. Therefore, our results support an explanation that focuses on a biased perception of group norms.

Lastly, our research also makes a methodological contribution to how to create group identification in experiments. In their research, Chen and Li (2009) examined how various manipulations influenced group identification. They found that including a chat stage (binary) significantly increased group identification by 0.126 (rescaled to a scale from 0 to 1). To compare their results with our results, we estimated an OLS regression model. In this model, we regressed the same group identification measure that Chen and Li used (see Appendix D.9) against the subject’s progressivism and the absolute difference between the individual’s progressivism and the social group’s implied progressivism. We find that increasing the distance progressivism by 1 (on a scale from 0 to 10) significantly decreased group identification by 0.088 (on a scale from 0 to 1). Since the mean distance progressivism was 1.28, the average moral similarity in our experiment had a similar effect on group identification as would be obtained by including a chat stage between the individual and Group A.

8 Conclusion

In this study, we found that group identification has a significant positive impact on rule compliance. Due to our design, we isolated this effect as being due to changes in group norm sensitivity rather than changes in the group norms to comply with the rule. Furthermore, we found that moral similarity based on moral foundations theory’s progressivism scale (Clark et al., 2017) has a positive impact on group identification.

Our work advances several important strands of literature. We show that group identification can influence behavior via group norm sensitivity. With this, our study identifies one situational variable that determines when social identities are likely to matter. Moreover, we show that similarity in moral preferences, just like similarity in other preferences (Chen & Li, 2009), increases group identification. We also contribute to the literature in management on adherence to organizational rules and norms to show that values alignment between the employee and the organization is a viable (and implementable) strategy that can lead employees to adopt policies and guidelines even when financial incentives discourage their adoption. Finally, on a methodological note, we show how one can create groups in the laboratory based on moral values and how the rule-following task (Kimbrough & Vostroknutov, 2018) can be modified to measure group norm sensitivity rather than general norm sensitivity. In principle, this method can measure the group norm

sensitivity towards any communicable social group.⁴¹

Our work identifies several promising avenues for future research. Our results suggest that it might be possible to describe and predict how an individual’s set of acquired social identities changes over time. At any point in time, an individual’s set of acquired social identities is (partly) determined by the individual’s and the available social groups’ moral positions. Hereby, previously acquired social identities determine an individual’s moral position⁴², and its members’ moral position determines a social group’s moral position. These connections might prove useful in the creation of dynamic social identity models. Second, it would be fruitful to study how extreme de-identification (despising another group) affects group norm sensitivity. It might be that individuals’ group norm sensitivity becomes negative, and individuals are willing to incur a cost to violate group norms. Finally, it would be valuable to explore whether similarity in different concepts of morality or specific constructs related to social preferences affects group identification similarly.

⁴¹It might even be used to measure “group norm sensitivity” towards atomistic social groups that only consist of one individual.

⁴²According to Haidt (2012), “*Moral systems are [...] interlocking sets of values, virtues, norms, practices, identities, institutions, technologies, and evolved psychological mechanisms that work together to suppress or regulate self-interest and make cooperative societies possible*”.

References

- Akerlof, G. A., & Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3), 715–753. doi:[10/fxdw82](#)
- Akerlof, G. A., & Kranton, R. E. (2002). Identity and schooling: Some lessons for the economics of education. *Journal of Economic Literature*, 40(4), 1167–1201. doi:[10/gd2x2h](#)
- Akerlof, G. A., & Kranton, R. E. (2005). Identity and the economics of organizations. *Journal of Economic Perspectives*, 19(1), 9–32. doi:[10/dkjmnm](#)
- Akerlof, G. A., & Kranton, R. E. (2008). Identity, supervision, and work groups. *American Economic Review*, 98(2), 212–217. doi:[10/dzq452](#)
- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, 63(4), 596–612. doi:[10/ddr983](#)
- Benjamin, D. J., Choi, J. J., & Strickland, A. J. (2010). Social identity and preferences. *American Economic Review*, 100(4), 1913–1928. doi:[10/fw6f2c](#)
- Bernard, M., Hett, F., & Mechtel, M. (2016). Social identity and social free-riding. *European Economic Review*, 90, 4–17. doi:[f9jcqn](#)
- Bernhard, H., Fehr, E., & Fischbacher, U. (2006). Group affiliation and altruistic norm enforcement. *American Economic Review*, 96(2), 217–221. doi:[10/bfnvc9](#)
- Bicchieri, C., Dimant, E., Gächter, S., & Nosenzo, D. (2022). Social proximity and the erosion of norm compliance. *Games and Economic Behavior*, 132, 59–72. doi:<https://doi.org/gqd3z2>
- Chang, D., Chen, R., & Krupka, E. (2019). Rhetoric matters: A social norms explanation for the anomaly of framing. *Games and Economic Behavior*, 116, 158–178. doi:[10/gm4dc2](#)
- Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9, 88–97. doi:[10/bj42](#)
- Chen, R., & Chen, Y. (2011). The potential of social identity for equilibrium selection. *American Economic Review*, 101(6), 2562–2589. doi:[10/chd9nt](#)
- Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, 99(1), 431–457. doi:[10/cm3z39](#)
- Clark, C. B., Swails, J. A., Pontinen, H. M., Bowerman, S. E., Kriz, K. A., & Hendricks, P. S. (2017). A behavioral economic assessment of individualizing versus binding moral foundations. *Personality and Individual Differences*, 112, 49–54. doi:[10/gfxk2c](#)
- D’Adda, G., Drouvelis, M., & Nosenzo, D. (2016). Norm elicitation in within-subject designs: Testing for order effects. *Journal of Behavioral and Experimental Economics*, 62, 1–7. doi:[10/gm4dc7](#)
- Dehghani, M., Johnson, K., Hoover, J., Sagi, E., Garten, J., Parmar, N. J., ... Graham, J. (2016). Purity homophily in social networks. *Journal of Experimental Psychology: General*, 145(3), 366–375. doi:[10/f8bqq8](#)

- Dimant, E. (2024). Hate trumps love: The impact of political polarization on social preferences. *Management Science*, 70(1), 1–31. doi:<https://doi.org/g5m6v8>
- Douglas, B. D., Ewell, P. J., & Brauer, M. (2023). Data quality in online human-subjects research: Comparisons between mturk, prolific, cloudresearch, qualtrics, and sona. *PLOS ONE*, 18(3), 1–17. doi:<https://doi.org/grx5s2>
- Eckel, C. C., & Grossman, P. J. (2005). Managing diversity by creating team identity. *Journal of Economic Behavior & Organization*, 58(3), 371–392. doi:[10/dj4wf8](https://doi.org/10/dj4wf8)
- Enke, B. (2020). Moral values and voting. *Journal of Political Economy*, 128(10), 3679–3729. doi:[10/gh7pgg](https://doi.org/10/gh7pgg)
- Erkut, H., Nosenzo, D., & Sefton, M. (2015). Identifying social norms using coordination games: Spectators vs. stakeholders. *Economics Letters*, 130, 28–31. doi:[10/f7ch6q](https://doi.org/10/f7ch6q)
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. doi:[10/bxjdcg](https://doi.org/10/bxjdcg)
- Franks, A. S., & Scherr, K. C. (2015). Using moral foundations to predict voting behavior: Regression models from the 2012 U.S. presidential election. *Analyses of Social Issues and Public Policy*, 15(1), 213–232. doi:[10/f7489z](https://doi.org/10/f7489z)
- Gächter, S., Gerhards, L., & Nosenzo, D. (2017). The importance of peers for compliance with norms of fair sharing. *European Economic Review*, 97, 72–86. doi:[10/gbrzmt](https://doi.org/10/gbrzmt)
- Gächter, S., Nosenzo, D., & Sefton, M. (2013). Peer effects in pro-social behavior: Social norms or social preferences? *Journal of the European Economic Association*, 11(3), 548–573. doi:[10/f4z2qb](https://doi.org/10/f4z2qb)
- Gächter, S., Starmer, C., & Tufano, F. (2015). Measuring the closeness of relationships: A comprehensive evaluation of the 'inclusion of the other in the self' scale. *PLOS ONE*, 10(6), 1–19. doi:[10/cx3d](https://doi.org/10/cx3d)
- Gächter, S., Starmer, C., & Tufano, F. (2017, June). *Revealing the economic consequences of group cohesion* (IZA Discussion Paper No. 10824). IZA. Bonn, Germany. Retrieved from <https://bit.ly/3nwsoe7>
- Goette, L., Huffman, D., & Meier, S. (2006). The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups. *American Economic Review*, 96(2), 212–216. doi:[10/cw7355](https://doi.org/10/cw7355)
- Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral Foundations Theory: The pragmatic validity of moral pluralism. In P. Devine & A. Plant (Eds.), *Advances in experimental social psychology* (Chap. 2, Vol. 47, pp. 55–130). doi:[10/gdcfn9](https://doi.org/10/gdcfn9)
- Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, 101(2), 366–385. doi:[10/cq64hc](https://doi.org/10/cq64hc)
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York: Vintage.
- Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, 133(4), 55–66. doi:[10/dnwzr8](https://doi.org/10/dnwzr8)

- Harris, D., Herrmann, B., Kontoleon, A., & Newton, J. (2015). Is it a norm to favour your own group? *Experimental Economics*, 18(3), 491–521. doi:[10/f7nvfz](https://doi.org/10.1007/s10683-024-09821-5)
- Hornsey, M. J. (2008). Social identity theory and self-categorization theory: A historical review. *Social and Personality Psychology Compass*, 2(1), 204–222. doi:[10/fshg98](https://doi.org/10.1111/j.1751-9004.2007.00098.x)
- Iyer, R., Koleva, S., Graham, J., Ditto, P., & Haidt, J. (2012). Understanding libertarian morality: The psychological dispositions of self-identified libertarians. *PLOS ONE*, 7(8), 1–23. doi:[10/f36mch](https://doi.org/10.1371/journal.pone.0043388)
- Jiang, J., & Li, S. X. (2019). Group identity and partnership. *Journal of Economic Behavior & Organization*, 160, 202–213. doi:[10/gm4tjw](https://doi.org/10.1016/j.jebo.2018.10.014)
- Johnson, K. M., Iyer, R., Wojcik, S. P., Vaisey, S., Miles, A., Chu, V., & Graham, J. (2014). Ideology-specific patterns of moral indifference predict intentions not to vote. *Analyses of Social Issues and Public Policy*, 14(1), 61–77. doi:[10/f6t6t6](https://doi.org/10.1215/00141801-1290000)
- Kimbrough, E. O., Krupka, E. L., Kumar, R., Murray, J. M., Ramalingam, A., Sánchez-Franco, S., ... Hunter, R. F. (2024). On the stability of norms and norm-following propensity: A cross-cultural panel study with adolescents. *Experimental Economics*, 27(2), 351–378. doi:[10.1007/s10683-024-09821-5](https://doi.org/10.1007/s10683-024-09821-5)
- Kimbrough, E. O., & Vostroknutov, A. (2016). Norms make preferences social. *Journal of the European Economic Association*, 14(3), 608–638. doi:[10/f8wj56](https://doi.org/10.1111/jeea.12256)
- Kimbrough, E. O., & Vostroknutov, A. (2018). A portable method of eliciting respect for social norms. *Economics Letters*, 168, 147–150. doi:[10/gdxgtc](https://doi.org/10.1016/j.econlet.2018.03.014)
- Kimbrough, E. O., & Vostroknutov, A. (2020, April). *A theory of injunctive norms*. SSRN. doi:[10/gm4dc8](https://doi.org/10.2139/ssrn.3544448)
- Kranton, R., Pease, M., Sanders, S., & Huettel, S. (2020). Deconstructing bias in social preferences reveals groupy and not-groupy behavior. *Proceedings of the National Academy of Sciences*, 117(35), 21185–21193. doi:[10/gg8h75](https://doi.org/10.1073/pnas.2011193117)
- Kranton, R. E., & Sanders, S. G. (2017). Groupy versus non-groupy social preferences: Personality, region, and political party. *American Economic Review*, 107(5), 65–69. doi:[10/gkq6kg](https://doi.org/10.3386/w23061)
- Krupka, E., Hoover, H., Eckel, C., Rosenblat, T., Ojumu, O., & Wilson, R. K. (2023). Multiple social identities cloud norm perception: Responses to covid-19 among university aged republicans and democrats. *Frontiers in Behavioral Economics*, 2. doi:[https://doi.org/nknx](https://doi.org/10.3389/fbeh.2023.1054448)
- Krupka, E. L., Leider, S., & Jiang, M. (2017). A meeting of the minds: Informal agreements and social norms. *Management Science*, 63(6), 1708–1729. doi:[10/gfxkz4](https://doi.org/10.1287/mnsc.2017.2682)
- Krupka, E. L., & Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, 11(3), 495–524. doi:[10/f42c9g](https://doi.org/10.1111/jeea.12048)
- Krysowski, E., & Tremewan, J. (2021). Why does anonymity make us misbehave: Different norms or less compliance? *Economic Inquiry*, 59(2), 776–789. doi:[10/gj8m7b](https://doi.org/10.1111/ein.12448)
- Lazear, E. P., & Shaw, K. L. (2007). Personnel economics: The economist’s view of human resources. *Journal of Economic Perspectives*, 21(4), 91–114. doi:[fhdq8g](https://doi.org/10.1215/00141801-2007-004)

- Le, B., Moss, W. B., & Mashek, D. (2007). Assessing relationship closeness online: Moving from an interval-scaled to continuous measure of including others in the self. *Social Science Computer Review*, 25(3), 405–409. doi:[10/fs4xh9](#)
- Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, 93(2), 234–249. doi:[10/czjxx2](#)
- Milesi, P. (2017). Moral foundations and voting intention in Italy. *Europe's Journal of Psychology*, 13(4), 667–687. doi:[10/gm4ddb](#)
- Palan, S., & Schitter, C. (2018). Prolific.ac—a subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17, 22–27. doi:[10/gftct9](#)
- Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, 153–163. doi:[10/f93997](#)
- Pickup, M., Kimbrough, E. O., & de Rooij, E. A. (2021). Expressive politics as (costly) norm following. *Political Behavior*. doi:[g8pz](#)
- Pryor, C., Perfors, A., & Howe, P. D. L. (2019). Even arbitrary norms influence moral decision-making. *Nature Human Behaviour*, 3(1), 57–62. doi:[10/gfrf2b](#)
- Schubert, T. W., & Otten, S. (2002). Overlap of self, ingroup, and outgroup: Pictorial measures of self-categorization. *Self and Identity*, 1(4), 353–376. doi:[10/dv3jnp](#)
- Shang, J., Reed, A., II, & Croson, R. (2008). Identity congruency effects on donations. *Journal of Marketing Research*, 45(3), 351–361. doi:[10/drj8dd](#)
- Shayo, M. (2009). A model of social identity with an application to political economy: Nation, class, and redistribution. *American Political Science Review*, 103(2), 147–174. doi:[ds97pq](#)
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks/Cole Publishing Company.
- Tropp, L. R., & Wright, S. C. (2001). Ingroup identification as the inclusion of ingroup in the self. *Personality and Social Psychology Bulletin*, 27(5), 585–600. doi:[10/cfmm5m](#)
- Turner, J. C. (1991). *Social influence*. Pacific Grove, USA: Thomson Brooks/Cole.
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: Self-categorization theory*. Oxford, UK: Basil Blackwell.
- Turner, J. C., & Reynolds, K. J. (2012). Self-categorization theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (Chap. 46, Vol. 2, pp. 399–417). doi:[10/gm4ddc](#)

- Van de Vyver, J., Houston, D. M., Abrams, D., & Vasiljevic, M. (2016). Boosting belligerence: How the July 7, 2005, London bombings affected liberals' moral foundations and prejudice. *Psychological Science*, 27(2), 169–177. doi:[10/gg6hfj](https://doi.org/10/gg6hfj)
- Veselý, Š. (2015). Elicitation of normative and fairness judgments: Do incentives matter? *Judgment and Decision Making*, 10(2), 191–197.
- Winterich, K. P., Zhang, Y., & Mittal, V. (2012). How political identity and charity positioning increase donations: Insights from moral foundations theory. *International Journal of Research in Marketing*, 29(4), 346–354. doi:[10/f4kzmm](https://doi.org/10/f4kzmm)

Appendix A Deviations from the Pre-Analysis Plan

We preregistered our analysis at the American Economic Association’s registry for randomized controlled trials (<https://doi.org/g5qw>). We deviated from the Pre-Analysis Plan in the following ways:

1. We renamed the behavioral experiment the choice experiment.
2. We renamed the communal index the binding index. This change also applied to words related to the index (e.g., communal treatment was changed to binding treatment).
3. For reasons of clarity, we restated some of our hypotheses without changing their contents.⁴³
4. To improve readability, we moved the analysis of the original Hypothesis 2 (“According to the measured group norms, the social appropriateness is linearly increasing in the number of balls placed in the norm compliant bucket”) to the appendix (see Appendix B.2).
5. While examining the original Hypothesis 2, we used a Tobit regression model instead of the preregistered OLS regression model. We did this to take into account the censored nature of the social appropriateness ratings. However, we can report that the coefficient on the number of balls in the rule-compliant bucket was positive and significant in the OLS regression model ($b = 0.033$, $p < 0.001$).
6. While examining Hypothesis 2, we only reported the result of a preregistered one-sided t-test. We did not report the results of a one-sided Wilcoxon-Mann-Whitney test. This is because a one-sided Wilcoxon-Mann-Whitney test requires that the shapes of the distributions in the two samples are identical. This assumption might have been unrealistic in our case. However, a one-sided Wilcoxon-Mann-Whitney test would have been significant ($p < 0.001$).
7. We renamed Hypothesis 4a and 4b into Hypothesis 3 and 4.
8. While examining Hypothesis 3, we only reported the results of a preregistered comparison of the number of balls in the rule-compliant bucket between treatments. We did not report the results of a conditional (fixed-effect) logit model (McFadden, 1974). We deviated from the Pre-Analysis plan because the correlation between two explanatory variables made a reliable estimation of coefficients impossible. Since both group norms were identical, an action’s monetary earnings and average social appropriateness were

⁴³For reference, the preregistered hypothesis were: “**Hypothesis 1 (Group Norms)**: According to individualizing subjects, the group norms present in the modified rule following task are identical in the individualizing and the communal treatment”, “**Hypothesis 2 (Group Norms)**: According to the measured group norms, the social appropriateness is linearly increasing in the number of balls placed in the norm compliant bucket”, “**Hypothesis 3 (Group Identification)**: Group identification is significantly higher for individualizing individuals that are assigned to the individualizing treatment compared to those that are assigned to the communal treatment.” and “**Hypothesis 4 (Norm Compliance) / Hypothesis 4a**: Compliance with the group norms is higher in the individualizing treatment compared to the communal treatment. / **Hypothesis 4b**: This treatment effect is mediated by changes in group identification.”

highly correlated.

9. While examining Hypothesis 3, we only reported the result of a preregistered one-sided t-test. For the same reason as described in the paragraph on Hypothesis 1, we did not report the results of a one-sided Wilcoxon-Mann-Whitney test. However, both tests provided the same results: If all observations were included, a one-sided Wilcoxon-Mann-Whitney test was insignificant ($p = 0.167$). If the observations assigned to the authority control question were excluded, a one-sided Wilcoxon-Mann-Whitney test was significant ($p = 0.030$).
10. While examining Hypothesis 4, we only reported the results of a Tobit regression model. For the same reason as described in the paragraph on Hypothesis 3, we did not report the results of a conditional (fixed-effect) logit model.
11. While examining Hypothesis 4, we did not report the results of the preregistered regression model. Instead, we reported the results of two models that either did not include the binding treatment dummy (model 1) or included additional control variables (model 2). The coefficient of group identification was also positive and significant in the preregistered regression model ($b = 14.821$, $p = 0.001$).
12. We did not report the results of all preregistered exploratory analyses that examine how robust our results are to variations in the norm compliance measure (e.g., more than 0 balls in the rule-compliant bucket). The reason for this omission was that group norms were linearly increasing in the number of balls in the rule-compliant bucket and did not abruptly change at some cutoff point (see appendix B.2).
13. We did not report the results of all preregistered exploratory analyses that examined the robustness of the relationship between moral similarity and group identification. We only reported those results of the Pre-Analysis Plan that did yield interesting results.

Appendix B Additional Analysis

B.1 Comparison of Moral and Socio-Demographic Data between Treatments

Table 6: Individual and Aggregated Moral Foundations Scores by Treatment and Experiment

	Choice experiment			Norm elicitation experiment		
	Individualizing treatment	Binding treatment	P-value ^a	Individualizing treatment	Binding treatment	P-value ^a
<i>Progressivism</i>	1.72	1.59	0.247	1.49	1.59	0.600
<i>Individualizing index</i>	3.89	3.83	0.596	3.82	3.85	0.716
<i>Binding index</i>	2.17	2.25	0.459	2.33	2.26	0.468
<i>Harm</i>	3.84	3.79	0.451	3.77	3.81	0.529
<i>Fairness</i>	3.94	3.88	0.409	3.88	3.88	0.957
<i>Ingroup</i>	2.38	2.52	0.292	2.62	2.54	0.634
<i>Authority</i>	2.14	2.10	0.886	2.15	2.18	0.820
<i>Purity</i>	1.98	2.13	0.315	2.22	2.06	0.246

Note. Progressivism could take on values from -5 to 5 . All other indices could take on values from 0 to 5 .

^a The p-values are based on a two-sided Wilcoxon-Mann-Whitney test that compared the scores between treatments.

This subsection compares the moral foundations and social-demographic data of participants of the choice and norm elicitation experiment between treatments. We found no differences between treatments. Consequently, we attributed any treatment effect to our treatment variation.

To examine this, we report in Table 6, separated by experiment and treatment, the average score of participants in the five individual moral foundations (harm, fairness, ingroup, authority, purity) and the three aggregated indices that we used (individualizing index, binding index, progressivism). Furthermore, in Tables 7 and 8 we report separated by experiment and treatment the average answer or the frequency distribution of answers to all socio-demographic questions (age, gender, ethnicity, degree, employment, income, community, religion, politics). We compared all moral and socio-demographic characteristics between the individualizing and binding treatment of the choice and norm elicitation experiment. In doing so, we used the appropriate test to either test for differences in mean (two-tailed Wilcoxon-Mann-Whitney test) or frequency distribution (Fisher’s exact test). To account for multiple testing (34 comparisons in total / 17 per experiment), we applied the Bonferroni correction to our statistical significance threshold of $p = 0.05$. The new threshold became $p = 0.001$. We did not detect any significant differences between treatments in the choice or norm elicitation experiment (in all cases $p > 0.053$).

Table 7: Answers to the Socio-Demographic Questions by Treatment and Experiment (Part 1)

	Choice experiment			Norm elicitation experiment		
	Individualizing treatment	Binding treatment	P-value ^a	Individualizing treatment	Binding treatment	P-value ^a
<i>Age</i>						
Mean	35.26	35.07	.	36.04	35.35	.
Total	159	160	0.947	99	100	0.323
<i>Gender</i>						
Female	87	87	.	53	54	.
Male	69	71	.	45	46	.
Other	3	2	.	1	0	.
Total	159	160	0.943	99	100	1.000
<i>Ethnicity</i>						
White	123	118	.	79	69	.
Asian	12	18	.	6	12	.
Hispanic	6	8	.	5	9	.
Black	14	10	.	5	6	.
Other	5	6	.	3	4	.
Total	160	160	0.681	98	100	0.405
<i>Degree</i>						
None	1	0	.	1	0	.
High school diploma	68	50	.	34	45	.
Bachelor degree	65	70	.	45	43	.
Master's degree	20	32	.	11	10	.
Doctorate degree	6	8	.	8	1	.
Total	160	160	0.111	99	99	0.067
<i>Employment</i>						
Full time employed	62	82	.	46	46	.
Part time employed	17	24	.	14	11	.
Unemployed	20	9	.	10	8	.
Self-employed	22	23	.	12	15	.
Student	19	11	.	9	16	.
Housewife / husband	8	5	.	3	2	.
Retired	11	6	.	5	2	.
Total	159	160	0.053	99	100	0.643

Note. Answering the Socio-Demographic Questions was voluntary.

^a The p-values are based on two-sided Wilcoxon-Mann-Whitney (age) or Fisher's exact tests (all other variables) that compared the answers between treatments.

Table 8: Answers to the Socio-Demographic Questions by Treatment and Experiment (Part 2)

	Choice experiment			Norm elicitation experiment		
	Individualizing treatment	Binding treatment	P-value ^a	Individualizing treatment	Binding treatment	P-value ^a
<i>Income</i>						
\$0 - \$10,000	40	25	.	18	23	.
\$10,000 - \$25,000	26	28	.	15	22	.
\$25,000 - \$50,000	39	44	.	30	24	.
\$50,000 - \$75,000	24	31	.	20	18	.
\$75,000 - \$100,000	11	14	.	4	7	.
\$100,000 - \$150,000	14	13	.	5	4	.
More than \$150,000	3	3	.	7	2	.
Total	157	158	0.520	99	100	0.392
<i>Community</i>						
Urban	44	55	.	29	33	.
Suburban	86	85	.	53	54	.
Rural	29	20	.	17	13	.
Total	159	160	0.243	99	100	0.700
<i>Religion</i>						
Christian	49	55	.	43	37	.
Atheist	57	50	.	31	28	.
Jew	1	3	.	3	2	.
Buddhist	3	2	.	2	2	.
Muslim	2	0	.	0	1	.
Other	47	49	.	20	29	.
Total	159	159	0.619	99	99	0.646
<i>Politics</i>						
Very liberal	44	38	.	19	18	.
Liberal	53	60	.	35	38	.
Moderate	45	52	.	29	33	.
Conservative	14	8	.	12	10	.
Very conservative	4	2	.	4	1	.
Total	160	160	0.462	99	100	0.708

Note. Answering the Socio-Demographic Questions was voluntary.

^a The p-values are based on two-sided Fisher's exact tests that compared the answers between treatments.

Table 9: The Effect of Rule Compliance on Social Appropriateness

	<i>Social appropriateness</i>			
	(1)		(2)	
	Coef.	Std. err.	Coef.	Std. err.
<i>Balls in the blue bucket</i>	0.048***	(0.007)	0.055***	(0.009)
<i>Binding treatment</i>			0.170	(0.201)
<i>Interaction</i>			-0.014	(0.013)
Constant	-0.373***	(0.080)	-0.531***	(0.150)
Controls	No		Yes	
Log-likelihood	-1,213		-1,202	
BIC	2,446		2,521	
Number of cases	1,000		1,000	

Notes. We used Tobit regressions with standard errors clustered at the level of the individual. The controls include the following variables: The subject's progressivism, the distance in progressivism, one dummy for each control question (except for the harm control question), interactions between the control question dummies and the binding treatment dummy, and the time spent on stages one and two. Coef., Coefficient; Std. err., Standard error.

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

B.2 Effect of the Number of Balls in the Blue Bucket on Social Appropriateness Ratings

As an intuition check, we also tested whether the social appropriateness of an action was linearly increasing with the number of balls placed in the rule-compliant bucket. Placing less than 20 balls into the bucket constituted partial rule compliance and should have been perceived as proportionally less socially appropriate. We found evidence that more balls in the blue bucket was rated as more socially appropriate.

We estimated two Tobit regression models with standard errors clustered at the individual level.⁴⁴ In both models, we used the social appropriateness of an action in the group rule-following task as our dependent variable. In the first model, we regressed the dependent variable against the number of balls placed in the rule-compliant blue bucket. In the second model, we added a binding treatment dummy and an interaction term of the two other explanatory variables. We also included a set of control variables: The subject’s progressivism, the distance in progressivism, one dummy for each control question except for the harm control question, interactions between the control question dummies and the binding treatment dummy, and the time spent on stages one, and two. We find that the estimated coefficients in Table 9. As can be seen, the coefficient of the number of balls in the rule-compliant bucket was positive and highly significant in all specifications ($p < 0.001$). Furthermore, the insignificant interaction suggests that this relation did not differ by treatment.

B.3 Effect of Treatment on Alternative Group Identification Measures: Overlap and Survey Question

This subsection examines the robustness of the effect of treatment on group identification. In Hypothesis 2, we conceptualized group identification with the distance measure of the “Inclusion of Ingroup in the Self Scale.” Using two other group identification measures we replicated the initial analysis. We did not find any noticeable difference due to changes in the conceptualization of group identification.

Instead of the distance measure of the “Inclusion of Ingroup in the Self Scale”, we could have used its overlap measure. This measure measures group identification as the percentage of **overlap** between the “You” and the “Group A” circle. The overlap identification measure can take on values ranging from 0 (weak identification / the circles do not overlap) to 1 (strong identification / The “You” circle is contained in the “Group A” circle).⁴⁵ We could have also used the identification question from the **questionnaire** at the

⁴⁴We chose Tobit regression models to handle the censoring at 0 and 1. However, the coefficients of “Balls in the blue bucket” would have been significant and positive in OLS regression models as well (available on request).

⁴⁵The literature of identity fusion (Jiménez et al., 2016; Swann, Gómez, Seyle, Morales, & Huici, 2009) assumes that an overlap between the circles symbolizes that personal and social image are aligned about the group attitude. Since participants in our experiments were introduced to an unknown group, this appeared to be unlikely. However, this measure presumably captures

end of our study (see Appendix D.9). In it, we asked participants to rate how closely attached they felt to Group A. The normalized question identification measure could take on values ranging from 0 (“Not at all attached”) to 1 (“very strongly attached”).⁴⁶

To test whether different measures of group identification are consequential to our results, we replicated our analysis from Hypothesis 2 with the measures mentioned above. According to the overlap identification measure, the average group identification was 0.28 in the individualizing and 0.07 in the binding treatment. The 0.21 point difference was significant using a one-sided t-test ($p < 0.001$). According to the survey questionnaire response, the average group identification was 0.54 in the individualizing and 0.34 in the binding treatment. The 0.20 point difference was significant using a one-sided t-test ($p < 0.001$).

B.4 Effect of the Authority Control Question on Rule Compliance

This subsection examines how randomly assigning a participant the authority control question influenced the behavior in the group rule-following task. We found that the authority control question increased the number of balls in the rule-compliant bucket in the binding treatment but not in the individualizing treatment.

In Hypothesis 3, we predicted that the number of balls in the rule-compliant bucket of the group rule-following task would be higher in the individualizing than the binding treatment. However, our statistical analysis only supported this hypothesis if we excluded observations randomly assigned the authority control question (66 out of 320 observations). The authority control question asked the participants to state how the social group answered the following question: “When you decide whether something is right or wrong, to what extent are the following considerations relevant to your thinking - Whether or not someone showed a lack of respect for authority?”. In the individualizing treatment, the correct answer was “Slightly relevant,” and in the binding treatment, “Somewhat relevant.” We found that this variation had a heterogeneous effect on our treatments. While assigning the authority control question did not influence the number of balls in the rule-compliant bucket in the individualizing treatment, it significantly increased the number of balls in the rule-compliant bucket of the binding treatment. This made it less likely to observe the treatment effect described in Hypothesis 3.

To examine how the authority control question influenced behavior in the group rule-following task, we plotted six histograms that show the number of balls in the rule-compliant bucket separated by treatment and by whether the participant was randomly assigned any control question, the authority control question, or any control question other than the authority control question (see Figure 4). In the individualizing treatment, assigning the authority control question had little effect on the behavior in the group rule-following

an extreme form of group identification.

⁴⁶The question was unincentivized and took place after the experiment. Consequently, while answering, participants might have desired to appear consistent with their behavior in the group rule-following task.

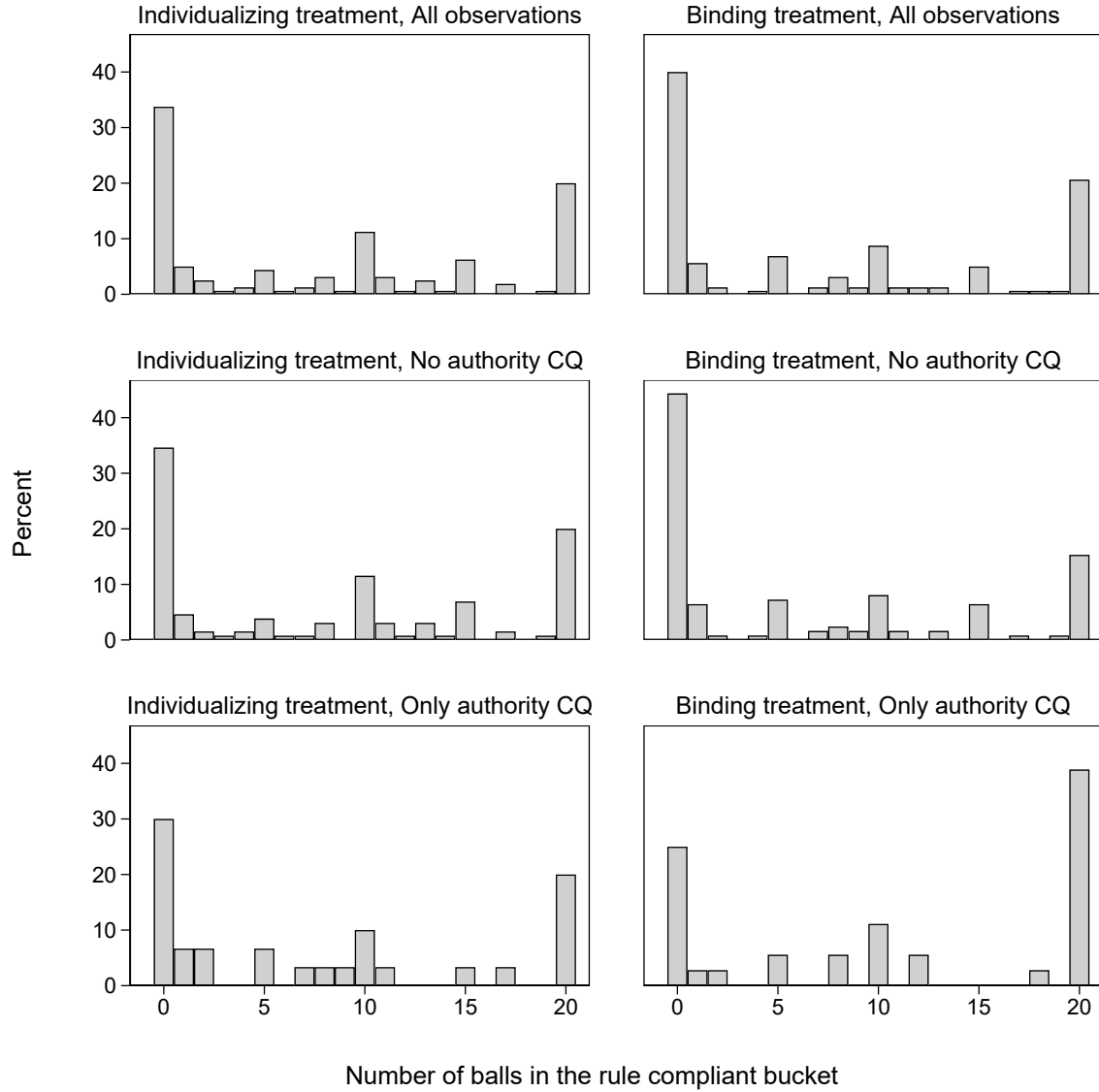


Figure 4: Number of Balls in the Rule-Compliant Bucket by Treatment and Control Question

Notes. For reference, 130 out of 160 observations in the individualizing treatment and 124 out of 160 observations in the binding treatment were not assigned the authority control question. CQ, Control Question.

task. However, in the binding treatment, assigning the authority control question reduced the number of participants that placed 0 balls in the rule-compliant bucket and increased the number of participants that placed 20 balls in the rule-compliant bucket.

To test this impression formally: Participants assigned to the individualizing treatment and any question other than the authority control question placed on average 8.24 balls in the rule-compliant bucket. Those individuals assigned to the individualizing treatment and the authority control question placed on average 7.77 balls in the rule-compliant bucket. According to a two-sided t-test, this difference of -0.47 balls in the rule-compliant bucket was not statistically significant ($p = 0.768$). On the other hand, participants assigned to the binding treatment and any question other than the authority control question placed on average 6.44 balls in the rule-compliant bucket. Those individuals assigned to the binding treatment and the authority control question placed on average 10.86 balls in the rule-compliant bucket. According to a two-sided t-test, this difference of 4.42 balls in the rule-compliant bucket was statistically significant ($p = 0.003$).

It appears that the authority control question increased the number of balls in the rule-compliant bucket in the binding treatment but not in the individualizing treatment. One reason for this might be that individuals were primed towards either weak or strong respect towards authority.⁴⁷ Consequently, individuals in the group rule-following task might have had a stronger desire to comply with arbitrary rules from an authority (Group A) in the binding treatment than the individualizing treatment.⁴⁸ We should note that the priming effect ran counter to the group identification effect we expected to see between treatments.

B.5 Effect of Alternative Group Identification Measures on Rule Compliance: Overlap and Survey Question

This subsection examines the robustness of the effect of group identification on rule compliance. In Hypothesis 4, we conceptualized group identification with the distance measure of the “Inclusion of Ingroup in the Self Scale.” Using two other group identification measures we replicated the initial analysis. We did not find any noticeable difference due to changes in the conceptualization of group identification.

To examine how group identification’s conceptualization influenced our results, we replicated our analysis from Hypothesis 4 with the **overlap** measure from the same scale or the identification question from the **questionnaire** at the end of our study (for more information, see Appendix B.3). We estimated four Tobit regression models: In all models, we used the number of balls in the rule-compliant bucket of the group rule-following task as our dependent variable. In the first model, we regressed the dependent variable against

⁴⁷A review of priming in economics can be found in Cohn and Maréchal (2016)

⁴⁸This finding contrasts the results of Kimbrough and Vostroknutov (2016), who did not observe any effect of the five moral foundations of moral foundations theory on rule-following in the traffic-light rule-following task (see Appendix E.1 of their study).

Table 10: The Effect of Alternative Group Identification Measures on Rule Compliance

	<i>Balls in the blue bucket</i>			
	(1)	(2)	(3)	(4)
<i>Identification (overlap)</i>	6.428*	7.161*		
	(2.954)	(3.139)		
<i>Identification (question)</i>			27.506***	30.995***
			(3.475)	(3.771)
<i>Binding treatment</i>		-6.153		-1.425
		(4.750)		(4.143)
Constant	4.258***	7.005	-6.648***	-9.625*
	(1.169)	(4.718)	(1.876)	(4.679)
Controls	No	Yes	No	Yes
Log-likelihood	-742	-729	-710	-694
BIC	1,500	1,556	1,436	1,486
Number of cases	320	320	320	320

Notes. We used Tobit regressions. The standard errors are displayed in parentheses. The controls include the following variables: The subject's progressivism, the distance in progressivism, one dummy for each control question (except for the harm control question), interactions between the control question dummies and the binding treatment dummy, and the time spent on stages one, two and three.

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

the overlap identification measure. In the second model, we added a binding treatment dummy and a set of control variables. The set included: The subject’s progressivism, the distance in progressivism, one dummy for each control question (except for the harm control question), interactions between the control question dummies and the binding treatment dummy, and the time spent on stages one, two, and three. In the third model, we regressed the dependent variable against the question identification measure. In the fourth model, we again added a binding treatment dummy and the set of control variables from model two. We report the results in Table 10. As can be seen, the identification measures were positive and significant in all four specifications ($p < 0.030$). However, changing the conceptualization of group identification changed the size of the estimated coefficient. If group identification increased by 1, the number of balls in the rule-compliant bucket, *ceteris paribus*, increased by approximately 7 (overlap measure), 14 (distance measure / see Table 5), or 29 (question measure). The model fit according to the Log-Likelihood, and the Bayesian Information Criterion increased in the same order. The reasons for this finding might have been that the overlap measure only captured an extreme form of group identification and that the question measure also captured the desire to appear consistent with previous choices in the group rule-following task. However, in general, it seems that our findings were robust to changes in the conceptualization of group identification.

B.6 Effect of Progressivism’s Sub-components on Group Identification: Individualizing Index and Binding Index, and Individual Moral Foundations

The progressivism index is a composite index. It describes the degree to which individuals endorse individualizing foundations (harm and fairness) over binding foundations (ingroup, authority, and purity). We replicated our analysis from the main section with the *sub-components* contained in the progressivism index. We found that the distance in the individualizing and the binding index both influenced group identification. Furthermore, this effect was mainly driven by the distance in the fairness foundation.

We first examined how distance in the individualizing and the binding index influenced group identification. To do so, we estimated two Tobit regression models with robust standard errors. In all models, we used the distance group identification measure of the “Inclusion of Ingroup in the Self Scale” as our dependent variable. In the first model, we regressed the dependent variable against the subject’s individualizing index, the subject’s binding index, the absolute difference between the individual’s individualizing index and the social group’s implied individualizing index, and the absolute difference between the individual’s binding index and the social group’s implied binding index. In the second model, we added a binding treatment dummy and a set of control variables. The set included: One dummy for each control question (except for the harm control question), interactions between the control question dummies and the binding treatment dummy, and the time spent on stages one and two. We report the results in Table 11. As can be seen, the coefficients for the distance in the individualizing and the binding index were negative and significant in both models ($p < 0.003$). This result suggests that both the similarity in the individualizing and the binding

Table 11: Decomposition of the Effect of Moral Similarity on Group Identification (Individualizing and Binding Indices)

	<i>Identification (distance)</i>			
	(1)		(2)	
	Coef.	Std. err.	Coef.	Std. err.
<i>Individualizing index</i>	0.098**	(0.030)	0.076*	(0.030)
<i>Binding index</i>	0.028	(0.017)	0.034*	(0.017)
<i>Distance individualizing index</i>	−0.185***	(0.032)	−0.112**	(0.038)
<i>Distance binding index</i>	−0.098***	(0.025)	−0.084***	(0.024)
<i>Binding treatment</i>			−0.191**	(0.066)
Constant	0.486***	(0.119)	0.578***	(0.125)
Controls	No		Yes	
Log-likelihood	-33		-19	
BIC	100		136	
Number of cases	320		320	

Notes. We used Tobit regressions with robust standard errors. The controls include the following variables: One dummy for each control question (except for the harm control question), interactions between the control question dummies and the binding treatment dummy, and the time spent on stages one and two. We report standard errors in the parentheses. Coef., Coefficient; Std. err., Standard error.

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

Table 12: Decomposition of the Effect of Moral Similarity on Group Identification
(Harm, Fairness, Ingroup, Authority, and Purity)

	<i>Identification (distance)</i>			
	(1)		(2)	
	Coef.	Std. err.	Coef.	Std. err.
<i>Harm</i>	0.061*	(0.029)	0.051	(0.028)
<i>Fairness</i>	0.046	(0.029)	0.029	(0.026)
<i>Ingroup</i>	0.003	(0.025)	−0.004	(0.025)
<i>Authority</i>	0.024	(0.027)	0.035	(0.026)
<i>Purity</i>	−0.011	(0.020)	−0.000	(0.018)
<i>Distance harm</i>	−0.023	(0.036)	−0.005	(0.036)
<i>Distance fairness</i>	−0.165***	(0.035)	−0.108**	(0.036)
<i>Distance ingroup</i>	−0.028	(0.023)	−0.036	(0.023)
<i>Distance authority</i>	−0.034	(0.032)	−0.039	(0.032)
<i>Distance purity</i>	−0.040	(0.026)	−0.014	(0.024)
<i>Binding treatment</i>			−0.196**	(0.063)
Constant	0.496***	(0.119)	0.589***	(0.123)
Controls	No		Yes	
Log-likelihood	-32		-17	
BIC	133		166	
Number of cases	320		320	

Notes. We used Tobit regressions with robust standard errors. The controls include the following variables: One dummy for each control question (except for the harm control question), interactions between the control question dummies and the binding treatment dummy, and the time spent on stages one and two. Coef., Coefficient; Std. err., Standard error.

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

moral values affected group identification.

To examine how the distance in the individual moral foundations influenced group identification, we estimated two additional Tobit regression models with robust standard errors. In all models, we used the distance group identification measure of the “Inclusion of Ingroup in the Self Scale” as our dependent variable. In the first model, we regressed the dependent variable against the subject’s scores in each of the five moral foundations (harm, fairness, authority, ingroup, and purity) and the absolute difference between the individual’s score and the implied social group’s score in each of the five moral foundations. In the second model, we added a binding treatment dummy and the same set of control variables as in model two of the first decomposition. We report the results in Table 12. As can be seen, only the distance in fairness was negative and significant ($p < 0.003$).⁴⁹

B.7 Effect of Alternative Moral Similarity Indices on Group Identification: Euclidean Distance and Chebyshev Distance

This subsection examines the robustness of the effect of moral similarity on group identification by using different aggregation methods than our progressivism index. We found that the effect of moral similarity on group identification was robust to changes in the conceptualization of moral similarity.

In the study’s main section, we conceptualized morality with the five foundations from moral foundations theory (harm, fairness, ingroup, authority, and purity) and used the progressivism index. That index aggregates all five foundations into one index by subtracting the individualizing foundations (harm and fairness) from the binding foundations (ingroup, authority, and purity). To create one index of (inverse) moral similarity, we calculated the absolute difference between the individual’s progressivism and the social group’s implied progressivism. However, we could have also calculated different moral similarity indices based on moral foundations theory. We could have calculated the Euclidean⁵⁰ and Chebyshev⁵¹ distance between the individual’s moral position and the social group’s moral position.

To examine how using these distance measures would have influenced our results, we estimated four Tobit regression models with robust standard errors. In all models, we used the distance group identification

⁴⁹Dehghani et al. (2016) used the moral foundations dictionary to measure the moral foundations of Twitter users as implied by their Tweets. They found that the distance in the purity, but not in any other moral foundation, predicted the distance between two people in the social network. Using lab studies, they showed that the distance in purity between two subjects predicted the preferred self-reported social and physical distance. Contrary to these results, we did not observe any effect of the distance in purity on group identification.

⁵⁰Euclidean distance puts equal weight on the distance of all five moral foundations. Unlike progressivism, Euclidean distance does not impose a distinction between individualizing and binding moral values.

⁵¹Chebyshev distance only considers the moral foundation that differs the most between the individual and the social group. Therefore it only takes into account the most salient difference between the individual and the social group.

Table 13: The Effect of Alternative Moral Similarity Metrics on Group Identification

	<i>Identification (distance)</i>			
	(1)	(2)	(3)	(4)
<i>Progressivism</i>	0.006 (0.016)	-0.003 (0.016)	-0.004 (0.016)	-0.011 (0.016)
<i>Euclidean distance</i>	-0.091*** (0.019)	-0.069*** (0.018)		
<i>Chebyshev distance</i>			-0.094*** (0.028)	-0.088*** (0.025)
<i>Binding treatment</i>		-0.234*** (0.066)		-0.258*** (0.064)
Constant	0.913*** (0.047)	0.974*** (0.066)	0.874*** (0.051)	0.978*** (0.070)
Controls	No	Yes	No	Yes
Log-likelihood	-49	-26	-56	-28
BIC	121	139	134	142
Number of cases	320	320	320	320

Notes. We used Tobit regressions with robust standard errors. The standard errors are displayed in parentheses. The controls include the following variables: One dummy for each control question (except for the harm control question), interactions between the control question dummies and the binding treatment dummy, and the time spent on stages one and two.

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

measure of the “Inclusion of Ingroup in the Self Scale” as our dependent variable. In the first model, we regressed the dependent variable against the subject’s progressivism and the Euclidean distance between the individual’s moral position and the social group’s moral position. In the second model, we added a binding treatment dummy and a set of control variables. The set included: One dummy for each control question (except for the harm control question), interactions between the control question dummies and the binding treatment dummy, and the time spent on stages one and two. In the third model, we regressed the dependent variable against the subject’s progressivism and the Chebyshev distance between the individual’s moral position and the social group’s moral position. In the fourth model, we added a binding treatment dummy and the same set of control variables as in model two. We report the results in Table 13. As can be seen, both the Euclidean distance and the Chebyshev distance were negative and highly significant in all models ($p < 0.001$). Furthermore, if one compares the log-likelihood and BIC of those models with the models reported in Table 5, one can see that the progressivism index had the highest goodness-of-fit out of all three measures.

Appendix C Additional Data Collection

To further investigate Result 3, we conducted additional data collections: the “limited control questions study” and the “no control questions study.” The additional data collection efforts allowed us (1) to test Hypothesis 3 without the priming effect of the authority control question and (2) to indirectly and imperfectly test whether specific awareness about moral similarity or dissimilarity with the rule-setting group impacts rule following.

In both additional data collections, we recruited new subjects for additional sessions of the prescreening survey and the choice experiment. The only major change in those additional data collections was how we introduced the social group in the choice experiment. In our main study, we showed participants five questions from the moral foundation questionnaire with the corresponding answers from the social group. Subsequently, we asked them to recall the social group’s answer to one randomly selected question. If they provided the correct response, they received a reward of £0.50. Otherwise, they received nothing. In the “limited control questions study”, we asked subjects to recall the answer to one randomly selected question but removed the authority question from the pool of possible questions they could see. In the “no control questions study”, we removed the incentivized control question entirely.

In both additional data collections, we collected 466 new observations (233 randomized into each treatment of the choice experiment).⁵² Both additional data collections were preregistered at the American Economic Association’s registry for randomized controlled trials (<https://doi.org/mcw3> and <https://doi.org/m675>).

In the following subsections, we present the results of the analyses of the additional data collections. Hereby, we replicated the analysis from the main study, with all observations from the “limited control questions study,” with those observations from the “limited control questions study,” which had a progressivism score above 1.17, and with all observations from the “no control questions study.” Note that we report the result of our regression models without control variables. The reason for this is that we only included the control variables in our analysis in the main paper to account for the effect of the authority control question.

The main results of all three analyses remained largely unchanged, with only minor variations observed in some variables. However, crucially, if we use all observations from the “limited control questions study,” we are able to confirm Hypothesis 3.

⁵²The number of participants was based on the following power calculation: We used all observations from the choice experiment of the main study who were not randomly assigned the authority control question to determine an effect size of 0.23. Subsequently, we used G-Power (Faul, Erdfelder, Lang, & Buchner, 2007) to calculate the minimum required sample size for a one-tailed t-test, an effect size of 0.23, an α -error probability of 0.05, a power of 0.80, and an even allocation between treatments. Based on this power calculation, we needed at least 233 subjects per treatment.

C.1 Results from the “limited control questions study”

In this subsection, we replicate the analysis from the main section with all observations from the “limited control questions study”. We found that, as predicted in Hypothesis 3, rule compliance in the group rule-following task of the choice treatment was significantly higher in the individualizing treatment.

- **Hypothesis 1:** We did not collect additional sessions of the norm elicitation experiment.
- **Hypothesis 2:** The average group identification was 0.84 and 0.71 in the individualizing and binding treatments, respectively. The 0.13 point difference was significant according to both a one-sided t-test ($p < 0.001$) and a Wilcoxon-Mann-Whitney test ($p < 0.001$).
- **Hypothesis 3:** The participants placed, on average, 9.41 and 7.79 balls in the rule-compliant bucket of the individualizing and the binding treatment, respectively. The 1.62 balls difference was significant according to both a one-sided t-test ($p = 0.012$) and a Wilcoxon-Mann-Whitney test ($p = 0.009$).
- **Hypothesis 4:** We re-estimated the first Tobit regression model. The coefficient of identification (distance) was positive and significant ($b = 8.689$, $p = 0.005$).
- **Chapter 6.5 The Effect of Moral Similarity on Group Identification:** We re-estimated the first Tobit regression model. The coefficient of progressivism was insignificant ($b = -0.011$, $p = 0.394$), whereas the coefficient of distance progressivism was negative and significant ($b = -0.085$, $p < 0.001$).
- **Appendix B.3 Effect of Treatment on Alternative Group Identification Measures: Overlap and Survey Question:** The average overlap identification measure was 0.34 and 0.17 in the individualizing and binding treatments, respectively. The 0.16 point difference was significant according to a one-sided t-test ($p < 0.001$). The average question identification measure was 0.54 and 0.39 in the individualizing and binding treatments, respectively. The 0.16 point difference was significant according to a one-sided t-test ($p < 0.001$).
- **Appendix B.5 Effect of Alternative Group Identification Measures on Rule Compliance: Overlap and Survey Question:** We re-estimated the first and third Tobit regression model. In the first model, the coefficient of identification (overlap) was positive and significant ($b = 6.147$, $p = 0.002$). In the third model, the coefficient of identification (question) was positive and significant ($b = 19.265$, $p < 0.001$).
- **Appendix B.6 Effect of Progressivism’s Sub-components on Group Identification: Individualizing Index and Binding Index, and Individual Moral Foundations:** We re-estimated the first Tobit regression model of both decompositions. In the first decomposition (into the individualizing index and binding index), the coefficients of the the following variables were insignificant: individualizing index ($b = 0.035$, $p = 0.169$). Whereas, the coefficients of the following variables were positive and significant: binding index ($b = 0.037$, $p = 0.008$). Furthermore, the coefficients of the following variables

were negative and significant: distance individualizing index ($b = -0.128$, $p < 0.001$) and distance binding index ($b = -0.081$, $p < 0.001$). In the second decomposition (into harm, fairness, authority, ingroup, and purity), the coefficients of the following variables were insignificant: harm ($b = -0.023$, $p = 0.063$), fairness ($b = 0.006$, $p = 0.629$), ingroup ($b = -0.001$, $p = 0.947$), authority ($b = 0.000$, $p = 0.963$), purity ($b = 0.014$, $p = 0.123$), distance ingroup ($b = -0.002$, $p = 0.890$), distance authority ($b = -0.026$, $p = 0.103$), and distance purity ($b = -0.012$, $p = 0.407$). Whereas the following variables were negative and significant: distance harm ($b = -0.036$, $p = 0.010$), and distance fairness ($b = -0.048$, $p = 0.001$).

- **Appendix B.7 Effect of Alternative Moral Similarity Indices on Group Identification: Euclidean Distance and Chebyshev Distance:** We re-estimated the first and third Tobit regression model. In the first model, the coefficient of progressivism was insignificant ($b = -0.023$, $p = 0.076$), whereas the coefficient of Euclidean distance was negative and significant ($b = -0.039$, $p < 0.001$). In the third model, the coefficient of progressivism was insignificant ($b = -0.027$, $p = 0.051$), whereas the coefficient of Chebyshev distance was negative and significant ($b = -0.041$, $p = 0.005$).

C.2 Results from the “limited control questions study”: Restricting Only To Observations With A Progressivism Score Above 1.17)

In this subsection, we replicate the analysis from the main section with observations from the “limited control questions study”, which had a progressivism score above 1.17.⁵³ We found that, as predicted in Hypothesis 3, rule compliance in the group rule-following task of the choice treatment was significantly higher in the individualizing treatment.

- **Hypothesis 1:** We did not collect additional sessions of the norm elicitation experiment.
- **Hypothesis 2:** The average group identification was 0.86 and 0.66 in the individualizing and binding treatments, respectively. The 0.20 point difference was significant according to both a one-sided t-test ($p < 0.001$) and a Wilcoxon-Mann-Whitney test ($p < 0.001$).

⁵³Both analyses for the “limited control questions study” were preregistered. The reason why we replicated the analysis from the main study, with all observations and only with those observations that had a progressivism score above 1.17, was the following: Participants that had a progressivism score between 0 and 1.17 were morally more similar to the binding rather than to the individualizing social group (implied progressivism score of the individualizing social group: 2.33 / implied progressivism score of the binding social group: 0). Consequently, those participants should have identified more strongly with the social group and placed more balls in the blue bucket of the rule-following task if assigned to the binding rather than the individualizing social group. Such a behavior would have counteracted the predicted treatment effect described in Hypothesis 2 and 3. Therefore, such behavior might explain why we could not confirm Hypothesis 3 in the main study. The reason why we used such an experimental design to test our hypothesis was that we initially assumed that the progressivism score of the binding group in the rule elicitation survey would be much lower and, thus, that participants would always be morally closer to the individualizing rather than the binding social group.

- **Hypothesis 3:** The participants placed, on average, 9.38 and 7.71 balls in the rule-compliant bucket of the individualizing and the binding treatment, respectively. The 1.67 balls difference was insignificant according to both a one-sided t-test ($p = 0.052$) and a Wilcoxon-Mann-Whitney test ($p = 0.071$). We attribute the lack of significance to the reduced number of observations.⁵⁴
- **Hypothesis 4:** We re-estimated the first Tobit regression model. The coefficient of identification (distance) was positive and significant ($b = 9.855$, $p = 0.026$).
- **Chapter 6.5 The Effect of Moral Similarity on Group Identification:** We re-estimated the first Tobit regression model. The coefficient of progressivism was insignificant ($b = -0.011$, $p = 0.693$), whereas the coefficient of distance progressivism was negative and significant ($b = -0.129$, $p < 0.001$).
- **Appendix B.3 Effect of Treatment on Alternative Group Identification Measures: Overlap and Survey Question:** The average overlap identification measure was 0.39 and 0.12 in the individualizing and binding treatments, respectively. The 0.27 point difference was significant according to a one-sided t-test ($p < 0.001$). The average question identification measure was 0.53 and 0.35 in the individualizing and binding treatments, respectively. The 0.19 point difference was significant according to a one-sided t-test ($p < 0.001$).
- **Appendix B.5 Effect of Alternative Group Identification Measures on Rule Compliance: Overlap and Survey Question:** We re-estimated the first and third Tobit regression model. In the first model, the coefficient of identification (overlap) was positive and significant ($b = 8.457$, $p = 0.006$). In the third model, the coefficient of identification (question) was positive and significant ($b = 24.111$, $p < 0.001$).
- **Appendix B.6 Effect of Progressivism's Sub-components on Group Identification: Individualizing Index and Binding Index, and Individual Moral Foundations:** We re-estimated the first Tobit regression model of both decompositions. In the first decomposition (into the individualizing index and binding index), the coefficients of the the following variables were insignificant: binding index ($b = 0.015$, $p = 0.632$). Whereas, the coefficients of the following variables were positive and significant: individualizing index ($b = 0.091$, $p = 0.018$). Furthermore, the coefficients of the following variables were negative and significant: distance individualizing index ($b = -0.183$, $p < 0.001$) and distance binding index ($b = -0.123$, $p < 0.001$). In the second decomposition (into harm, fairness, authority, ingroup, and purity), the coefficients of the following variables were insignificant: harm ($b = -0.018$, $p = 0.298$), fairness ($b = 0.029$, $p = 0.206$), ingroup ($b = 0.005$, $p = 0.756$), authority ($b = -0.021$, $p = 0.389$), purity

⁵⁴While excluding observations below the progressivism threshold had little effect on the effect size (all observations: 0.209 / above threshold: 0.208), it almost halved the number of usable observations per treatment (all observations: 233 / above threshold: 133). As a consequence, the power of the t-test decreased from 0.730 in the comparison with all observations to 0.493 in the comparison with only those observations above the progressivism threshold.

($b = 0.017$, $p = 0.131$), distance ingroup ($b = -0.019$, $p = 0.286$), and distance purity ($b = -0.013$, $p = 0.484$). Whereas the following variables were negative and significant: distance harm ($b = -0.051$, $p = 0.012$), distance fairness ($b = -0.073$, $p = 0.002$), and distance authority ($b = -0.071$, $p = 0.021$).

- **Appendix B.7 Effect of Alternative Moral Similarity Indices on Group Identification: Euclidean Distance and Chebyshev Distance:** We re-estimated the first and third Tobit regression model. In the first model, the coefficient of progressivism was insignificant ($b = -0.038$, $p = 0.171$), whereas the coefficient of Euclidean distance was negative and significant ($b = -0.079$, $p < 0.001$). In the third model, the coefficients of progressivism ($b = -0.056$, $p = 0.049$) and Chebyshev distance were negative and significant ($b = -0.091$, $p < 0.001$).

C.3 Results from the “no control questions study”

In this subsection, we replicate the analysis from the main section with all observations from the “no control questions study.” This additional data collection allows us to indirectly and imperfectly test whether specific awareness about moral similarity or dissimilarity with the rule-setting group impacts rule following.

- **Hypothesis 1:** We did not collect additional sessions of the norm elicitation experiment.
- **Hypothesis 2:** The average group identification was 0.80 and 0.67 in the individualizing and binding treatments, respectively. The 0.14 point difference was significant according to both a one-sided t-test ($p < 0.001$) and a Wilcoxon-Mann-Whitney test ($p < 0.001$).
- **Hypothesis 3:** The participants placed, on average, 7.74 and 7.08 balls in the rule-compliant bucket of the individualizing and the binding treatment, respectively. The 0.66 balls difference was insignificant according to a one-sided t-test ($p = 0.172$) and a Wilcoxon-Mann-Whitney test ($p = 0.188$). One reason for replicating the null result might be that the omission of the control question reduced the attention on the social groups’ moral positions and thus reduced the impact of our treatment variation.
- **Hypothesis 4:** We re-estimated the first Tobit regression model. The coefficient of identification (distance) was negative and significant ($b = 9.039$, $p = 0.002$).
- **Chapter 6.5 The Effect of Moral Similarity on Group Identification:** We re-estimated the first Tobit regression model. The coefficient of progressivism was insignificant ($b = -0.001$, $p = 0.935$), whereas the coefficient of distance progressivism was negative and significant ($b = -0.079$, $p < 0.001$).
- **Appendix B.3 Effect of Treatment on Alternative Group Identification Measures: Overlap and Survey Question:** The average overlap identification measure was 0.30 and 0.12 in the individualizing and binding treatments, respectively. The 0.18 point difference was significant according to a one-sided t-test ($p < 0.001$). The average question identification measure was 0.46 and 0.35 in the individualizing and binding treatments, respectively. The 0.11 point difference was significant according to a one-sided t-test ($p < 0.001$).

- **Appendix B.5 Effect of Alternative Group Identification Measures on Rule Compliance: Overlap and Survey Question:** We re-estimated the first and third Tobit regression model. In the first model, the coefficient of identification (overlap) was insignificant ($b = 3.527$, $p = 0.088$). In the third model, the coefficient of identification (question) was positive and significant ($b = 22.522$, $p < 0.001$).
- **Appendix B.6 Effect of Progressivism’s Sub-components on Group Identification: Individualizing Index and Binding Index, and Individual Moral Foundations:** We re-estimated the first Tobit regression model of both decompositions. In the first decomposition (into the individualizing index and binding index), the coefficients of the following variables were insignificant: individualizing index ($b = 0.043$, $p = 0.063$) and binding index ($b = 0.025$, $p = 0.076$). Whereas the following variables were negative and significant: distance individualizing index ($b = -0.090$, $p = 0.001$) and distance binding index ($b = -0.115$, $p < 0.001$). In the second decomposition (into harm, fairness, authority, ingroup, and purity), the coefficients of the following variables were insignificant: harm ($b = 0.043$, $p = 0.067$), fairness ($b = 0.003$, $p = 0.916$), ingroup ($b = 0.006$, $p = 0.752$), authority ($b = 0.012$, $p = 0.556$), purity ($b = 0.001$, $p = 0.929$), distance harm ($b = -0.029$, $p = 0.302$), and distance authority ($b = -0.042$, $p = 0.072$). Whereas the following variables were negative and significant: distance fairness ($b = -0.081$, $p = 0.006$), distance ingroup ($b = -0.046$, $p = 0.014$), and distance purity ($b = -0.045$, $p = 0.020$).
- **Appendix B.7 Effect of Alternative Moral Similarity Indices on Group Identification: Euclidean Distance and Chebyshev Distance:** We re-estimated the first and third Tobit regression model. In the first model, the coefficient of progressivism was insignificant ($b = -0.010$, $p = 0.443$), whereas the coefficient of Euclidean distance was negative and significant ($b = -0.086$, $p < 0.001$). In the third model, the coefficient of progressivism was insignificant ($b = -0.015$, $p = 0.268$), whereas the coefficient of Chebyshev distance was negative and significant ($b = -0.106$, $p < 0.001$).

Appendix D Experimental Materials

In the following subsections, we report the instructions of our online experiments. With this, it should be noted that: 1.) Everything below a screen tag (e.g., “**Screen 1:**”) was displayed on the same screen. 2.) Everything below a horizontal separation line was displayed in one text box. If the separation line includes centered text, the text was used as a text box title. 3.) Input items are always described by stating the used questions (“Q”) and all possible answers among which the participants could choose (“A”). 4.) In some cases we include comments (e.g., “[*Example comment*]”) that were not included in the online experiments. 5.) We do not include small details (e.g., buttons and their labels).

D.1 Moral Foundations Questionnaire

Screen 1:

Information:

In Part 1, we want you to answer a questionnaire. Please read the questions carefully and answer truthfully.

When you decide whether something is right or wrong, to what extent are the following considerations relevant to your thinking? Please rate each statement using this scale:

[5] = Extremely relevant (This is one of the most important factors when I judge right and wrong)

[4] = Very relevant

[3] = Somewhat relevant

[2] = Slightly relevant

[1] = Not very relevant

[0] = Not at all relevant (This consideration has nothing to do with my judgments of right and wrong)

Q1: “Whether or not someone suffered emotionally”

Q2: “Whether or not someone’s action showed love for his or her country”

Q3: “Whether or not someone showed a lack of respect for authority ”

Q4: “Whether or not someone violated standards of purity and decency”

Q5: “Whether or not someone was good at math”

Q6: “Whether or not some people were treated differently than others”

Q7: “Whether or not someone cared for someone weak or vulnerable”

Q8: “Whether or not someone acted unfairly”

Q9: “Whether or not someone did something to betray his or her group”

Q10: “Whether or not someone conformed to the traditions of society”

Q11: “Whether or not someone did something disgusting”

Q12: “Whether or not someone was cruel”

Q13: “Whether or not someone was denied his or her rights”

Q14: “Whether or not someone showed a lack of loyalty”

Q15: “Whether or not an action caused chaos or disorder”

Q16: “Whether or not someone acted in a way that God would approve of”

A: “Extremely relevant”, “Very relevant”, “Somewhat relevant”, “Slightly relevant”, “Not very relevant”
or “Not at all relevant”

Screen 2:

Information:

In Part 1, we want you to answer a questionnaire. Please read the questions carefully and answer truthfully.

Please read the following sentences and indicate your agreement or disagreement:

Q17: “Compassion for those who are suffering is the most crucial virtue”

Q18: “When the government makes laws, the number one principle should be ensuring that everyone is treated fairly”

Q19: “I am proud of my country’s history”

Q20: “Respect for authority is something all children need to learn”

Q21: “People should not do things that are disgusting, even if no one is harmed”

Q22: “It is better to do good than to do bad”

Q23: “One of the worst things a person could do is hurt a defenseless animal”

Q24: “Justice is the most important requirement for a society”

Q25: “People should be loyal to their family members, even when they have done something wrong”

Q26: “Men and women each have different roles to play in society”

Q27: “I would call some acts wrong on the grounds that they are unnatural”

Q28: “It can never be right to kill a human being”

Q29: “I think it’s morally wrong that rich children inherit a lot of money while poor children inherit nothing”

Q30: “It is more important to be a team player than to express oneself”

Q31: “If I were a soldier and disagreed with my commanding officer’s orders, I would obey anyway because that is my duty”

Q32: “Chastity is an important and valuable virtue”

A: “Strongly agree”, “Moderately agree”, “Slightly agree”, “Slightly disagree”, “Moderately disagree” or “Strongly disagree”

[We adopted the moral foundations questionnaire from Graham et al. (2011). The questionnaire consisted of two parts: The moral relevance section and the moral judgment section. We converted the categorical ratings of both sections into integers from 0 (“Not at all relevant” / “Strongly disagree”) to 5 (“Extremely relevant” / “Strongly agree”). The questions were either included to measure the harm foundation (Q1, Q7, Q12, Q17, Q23, Q28), the fairness foundation (Q2, Q8, Q13, Q18, Q24, Q29), the ingroup foundation (Q3, Q9, Q14, Q19, Q25, Q30), the authority foundation (Q4, Q10, Q15, Q20, Q26, Q31), the purity foundation (Q5, Q11, Q16, Q21, Q27, Q32) or as attention checks (Q6, Q22). The foundations were calculated by taking the mean of all contained items. In addition, we calculated the individualizing index by averaging the score of the harm and fairness foundations and the binding index by averaging the scores of the ingroup, authority and purity foundations. Lastly, we calculated the progressivism index by subtracting the individualizing index from the binding index (Clark et al., 2017). This index could take on values ranging from -5 (not at all individualizing & extremely binding) to $+5$ (extremely individualizing & not at all binding).]

D.2 Rule Elicitation Task

Screen 1:

Information:

In Part 1, we want you to read about a situation that other people will be facing in the future. While you yourself will **not** be in this situation, we want you to read about it, and then we will ask you a question. At the following stage, you will be able to reopen the text describing the situation.

Please take a moment to read about this situation:

In today’s study, you will decide how to allocate 20 balls between two buckets. Your task is to put each of the balls, one-by-one, into one of the two buckets: the blue bucket or the yellow bucket. The balls will appear in the center of your screen, and you can allocate each ball by clicking and dragging it to the bucket of your choice. For each ball you put in the blue bucket, you will receive £0.05, and for each ball you put in the yellow bucket, you will receive £0.10.

The rule is to put the balls in the blue bucket

Your payment will be based on your decisions: it is the sum of payments from the blue and yellow buckets.

Screen 2:

Information:

In Part 1, we want you to answer a question. Please read the question carefully and answer truthfully.

Q: “If you had to describe the rule to other study participants, how would you describe it?”

A: “The rule is to put the balls in the blue bucket” or “The rule is to put the balls in the yellow bucket”

D.3 Introduction to Group A

Screen 1:

Information:

In Part 1, we want to introduce you to another group of participants. During the rest of this study, we will call this group “**Group A**”. Some time ago, Group A took part in another study. Among other things, this study included a questionnaire. Below you can see the answers of Group A to a small selection of five questions. In each case, the answer is the rating that is closest to the average rating that we received from all members of Group A. We want you to carefully read and memorize the answers provided by members of Group A. On the next page, we will show you one question, and we will ask you to select the same response as members of Group A. If you choose the correct answer, you will receive a bonus of £0.50! We will tell you whether you were right at the end of the study.

When deciding whether something is right or wrong, to what extent are the following considerations relevant to the thinking of Group A? They rated each statement with the following scale:

[5] = Extremely relevant (This is one of the most important factors when I judge right and wrong)

[4] = Very relevant

[3] = Somewhat relevant

[2] = Slightly relevant

[1] = Not very relevant

[0] = Not at all relevant (This consideration has nothing to do with my judgments of right and wrong)

1. Whether or not someone cared for someone weak or vulnerable: [*In the individualizing treatment:* Very Relevant / *in the binding treatment:* Somewhat Relevant]
2. Whether or not some people were treated differently from others: [*In the individualizing treatment:* Very Relevant / *in the binding treatment:* Somewhat Relevant]
3. Whether or not someone's action showed love for his or her country: [*In the individualizing treatment:* Not very relevant / *in the binding treatment:* Somewhat Relevant]
4. Whether or not someone showed a lack of respect for authority: [*In the individualizing treatment:* Slightly relevant / *in the binding treatment:* Somewhat Relevant]
5. Whether or not someone violated standards of purity and decency: [*In the individualizing treatment:* Slightly relevant / *in the binding treatment:* Somewhat Relevant]

Screen 2:

Information:

In the box below, you can see one of the five questions presented on the last page. Please, select the rating that is closest to the average rating that we received from all members of Group A. If you choose the correct answer, you will receive a bonus of £0.50! We will tell you whether you were right at the end of the study.

When deciding whether something is right or wrong, to what extent is the following consideration relevant to the thinking of Group A? Please rate the statement in the same way as members of Group A did by using this scale:

[5] = Extremely relevant (This is one of the most important factors when they judge right and wrong)

[4] = Very relevant

[3] = Somewhat relevant

[2] = Slightly relevant

[1] = Not very relevant

[0] = Not at all relevant (This consideration has nothing to do with their judgments of right and wrong)

Q: “Whether or not someone cared for someone weak or vulnerable:”, “Whether or not some people were treated differently from others:”, “Whether or not someone’s action showed love for his or her country:”, “Whether or not someone showed a lack of respect for authority:” or “Whether or not someone violated standards of purity and decency:”

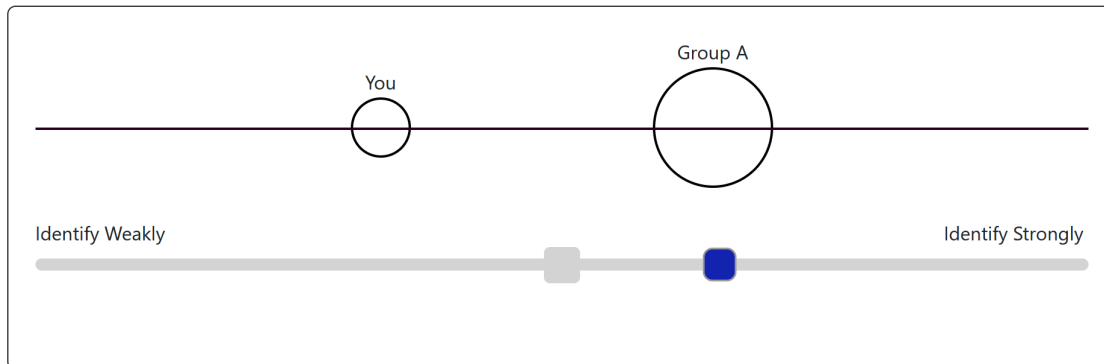
A: “Extremely relevant”, “Very relevant”, “Somewhat relevant”, “Slightly relevant”, “Not very relevant” or “Not at all relevant.”

D.4 Group Identification Task

Screen 1:

Information:

In Part 2, we want to know your identification with Group A. On the scale below you can see two circles: One representing you and one representing Group A. You can change the distance between the two circles by moving the blue slider on the scale. Moving the slider to the left moves the circles further apart and signals that you identify very little with Group A. Moving the slider to the right moves the circles closer to each other and signals that you identify very much with Group A. Please move the slider to a position in which the distance between the two circles most accurately describes your identification with Group A. Please zoom out if the scale is only displayed partially.



D.5 Rule Following Task

Screen 1:

Information:

In Part 3, you are going to play a game. You will receive the instructions for this game on the next page. Some parts of these instructions have been select by members of Group A. In addition to answering the questions that were presented to you in Part 1, the members of Group A also were asked to select a description for the game you are about to play. The description that was chosen by the majority was included in the instruction

you are about to receive.

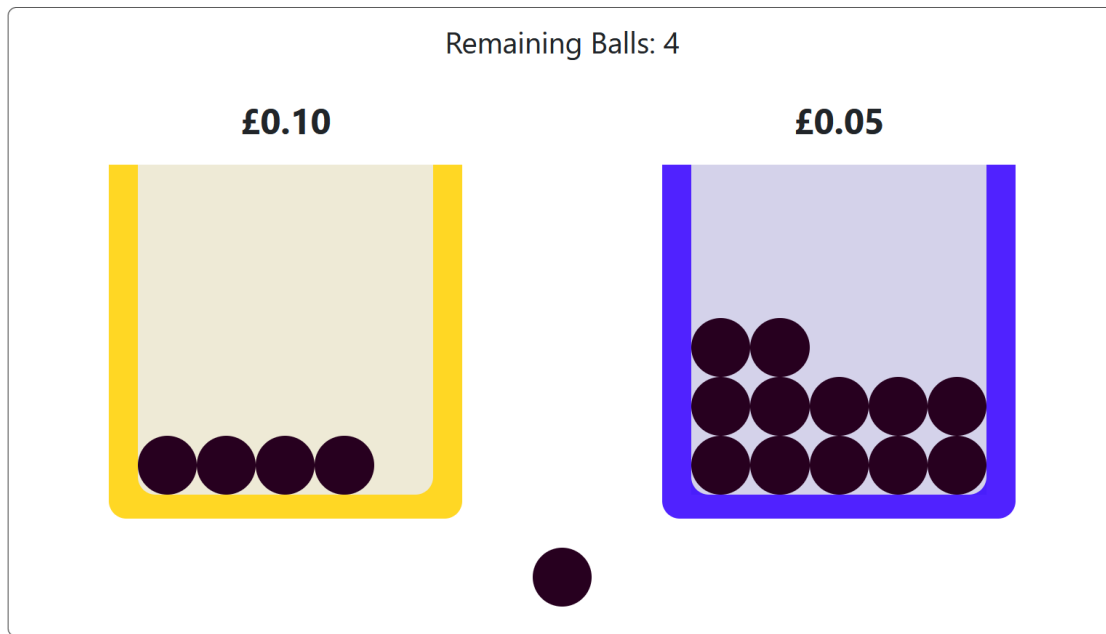
Screen 2:

Information:

In Part 3 of this study, you will decide how to allocate 20 balls between two buckets. Your task is to put each of the balls, one-by-one, into one of the two buckets: the blue bucket or the yellow bucket. The balls will appear in the center of your screen, and you can allocate each ball by clicking and dragging it to the bucket of your choice. For each ball you put in the blue bucket, you will receive £0.05, and for each ball you put in the yellow bucket, you will receive £0.10.

According to the members of Group A, the rule is to put the balls into the blue bucket

Your payment from Part 3 will be based on your decisions: it is the sum of payments from the blue and yellow buckets. After you have completed the task, a “Continue” button will be displayed at the bottom of the page.



D.6 Norm Elicitation

Screen 1:

Information:

In Part 2, we ask you to evaluate certain behaviors that may have occurred in another study. Up until this point, the participants of the other study had a very similar experience to you: Before participating

in the study, they answered the same questionnaire that you answered (more on this later), and they were introduced to the same Group A that you were introduced to in Part 1. We want you to read the instructions for one of their subsequent tasks. While you yourself will not play this game, you should read the instructions carefully. After you have read those instructions, you will receive further instructions that will detail your own task.

You will be able to reopen these instructions when you make your final decisions!

————— **Please take a moment to read about the task that you have to evaluate:** —————

Page 1:

In this part, you are going to play a game. You will receive the instructions for this game on the next page. Some parts of these instructions have been select by members of Group A. In addition to answering the questions that were presented to you in Part 1, the members of Group A also were asked to select a description for the game you are about to play. The description that was chosen by the majority was included in the instruction you are about to receive.

Page 2:

In today’s study, you will decide how to allocate 20 balls between two buckets. Your task is to put each of the balls, one-by-one, into one of the two buckets: the blue bucket or the yellow bucket. The balls will appear in the center of your screen, and you can allocate each ball by clicking and dragging it to the bucket of your choice. For each ball you put in the blue bucket, you will receive £0.05, and for each ball you put in the yellow bucket, you will receive £0.10.

According to the members of Group A, the rule is to put the balls into the blue bucket

Your payment will be based on your decisions: it is the sum of payments from the blue and yellow buckets.

Screen 2:

————— **Information:** —————

Your task in Part 2 is to evaluate how socially appropriate it is to put 0, 5, 10, 15, or 20 balls into the blue bucket while placing the remaining balls into the yellow bucket. In your judgment, you can rate any behavior on a scale that ranges from “Very socially inappropriate” to “Very socially appropriate”. Hereby, “Very socially inappropriate” indicates that the behavior is inconsistent with what most people expect an Individual ought to do, whereas “Very socially appropriate” indicates that the behavior was consistent with what most people expect an Individual ought to do. To further support you in your decision: A behavior might be considered as socially appropriate if most people agree that it is the “proper” thing to do, whereas a behavior might be considered as socially inappropriate if people might get angry if they observe a person doing it. You can find screenshots of your task and of your possible answers in the box below.

You will be able to reopen these instructions when you make your final decisions!

Screenshots of your upcoming task:

How socially appropriate is it ...

... to put 0 balls into the blue bucket and 20 balls into the yellow bucket:

... to put 5 balls into the blue bucket and 15 balls into the yellow bucket:

... to put 10 balls into the blue bucket and 10 balls into the yellow bucket:

... to put 15 balls into the blue bucket and 5 balls into the yellow bucket:

... to put 20 balls into the blue bucket and 0 balls into the yellow bucket:

How socially appropriate is it ...

... to put 0 balls into the blue bucket and 20 balls into the yellow bucket:

... to put 5 balls into the blue bucket and 15 balls into the yellow bucket:

... to put 10 balls into the blue bucket and 10 balls into the yellow bucket:

... to put 15 balls into the blue bucket and 5 balls into the yellow bucket:

... to put 20 balls into the blue bucket and 0 balls into the yellow bucket:

Very socially appropriate

Socially appropriate

Somewhat socially appropriate

Somewhat socially inappropriate

Socially inappropriate

Very socially inappropriate

Screen 3:

Information:

But wait, there's more! We don't want you to tell us whether you personally believe that the behavior is socially appropriate or inappropriate. Instead, we want you to guess how most other's from a specific group of people would rate the behavior. In the following, we will call this specific group of people, the **"Matching Group"**. You should note that the Matching Group and Group A consist of different people. We will tell you more about the Matching Group on the next page.

Learning about your Matching Group is important because your earnings will depend on it. At the end of the study, we will randomly select one of the five situations (put 0, 5, 10, 15, or 20 balls into the blue bucket) for all participants. Then we will compare your evaluation in this situation with the most frequent evaluation that we received from all members of your Matching Group (this includes you). If you selected the same appropriateness rating as was most frequent in your Matching Group, you will receive a bonus of £2.00. If

you selected a different rating, you will earn no bonus.

You will be able to reopen these instructions when you make your final decisions!

Screen 4:

Information:

Now we will explain to you who the members of your Matching Group are. Before participating in this study, each participant answered the same questionnaire. This questionnaire contained two types of questions (**category 1** and **category 2**). The questions included in the two categories differ in that they are concerned with different aspects of morality. For some people, the considerations expressed in category 1 are more important than those expressed in category 2, whereas for other people, the opposite is true. Lastly, some people might value both categories equally. Your answers indicate that:

For you, the considerations of category 1 are more important to those of category 2!

We call the group of participants, for whom the statements from category 1 were more relevant and agreeable than those of category 2, your **Matching Group**. Please read the questions contained in each category carefully and try to understand what both categories express. It might help you to imagine what kind of person would prefer either category. It is crucial that you understand who is in your Matching Group since it will help you to give us the same response as to what most people from your Matching Group believe constitutes socially appropriate or socially inappropriate behavior. Consequently, it will make it more likely for you to receive the additional £2.00.

You will be able to reopen these instructions when you make your final decisions!

Category 1:

To what extent are the following considerations relevant to your thinking about right and wrong:

- Whether or not someone suffered emotionally
- Whether or not some people were treated differently than others
- Whether or not someone cared for someone weak or vulnerable
- Whether or not someone acted unfairly
- Whether or not someone was cruel
- Whether or not someone was denied his or her rights

Indicate your agreement or disagreement:

- Compassion for those who are suffering is the most crucial virtue.

- When the government makes laws, the number one principle should be ensuring that everyone is treated fairly.
- One of the worst things a person could do is hurt a defenseless animal.
- Justice is the most important requirement for a society.
- It can never be right to kill a human being.
- I think it's morally wrong that rich children inherit a lot of money while poor children inherit nothing.

Category 2:

To what extent are the following considerations relevant to your thinking about right and wrong:

- Whether or not someone's action showed love for his or her country
- Whether or not someone showed a lack of respect for authority
- Whether or not someone violated standards of purity and decency
- Whether or not someone did something to betray his or her group
- Whether or not someone conformed to the traditions of society
- Whether or not someone did something disgusting
- Whether or not someone showed a lack of loyalty
- Whether or not an action caused chaos or disorder
- Whether or not someone acted in a way that God would approve of

Indicate your agreement or disagreement:

- I am proud of my country's history.
- Respect for authority is something all children need to learn.
- People should not do things that are disgusting, even if no one is harmed.
- People should be loyal to their family members, even when they have done something wrong.
- Men and women each have different roles to play in society.
- I would call some acts wrong on the grounds that they are unnatural.
- It is more important to be a team player than to express oneself.
- If I were a soldier and disagreed with my commanding officer's orders, I would obey anyway because that is my duty.

- Chastity is an important and valuable virtue.

Screen 5:

Information:

Please, rate now how socially appropriate it is to put 0, 5, 10, 15, or 20 balls into the blue bucket while placing the remaining balls into the yellow bucket. Should you have any questions remaining, you can reopen any of the previous instructions by clicking on the buttons below.

How socially appropriate is it ...

Q1: "... to put 0 balls into the blue bucket and 20 balls into the yellow bucket:"

Q2: "... to put 5 balls into the blue bucket and 15 balls into the yellow bucket:"

Q3: "... to put 10 balls into the blue bucket and 10 balls into the yellow bucket:"

Q4: "... to put 15 balls into the blue bucket and 5 balls into the yellow bucket:"

Q5: "... to put 20 balls into the blue bucket and 0 balls into the yellow bucket:"

A: "Very socially appropriate", "Socially appropriate", "Somewhat socially appropriate", "Somewhat socially inappropriate", "Socially inappropriate" or "Very socially inappropriate"

D.7 Socio-Demographic Questionnaire

Screen 1:

Information:

Before we inform you about your earnings and transfer you back to Prolific, we would like to ask you some optional questions. In doing so, please consider that all your data is collected anonymously and will not affect your payment!

1. [*Age*]

Q: "What is your age?"

A: Any integer number

2. [*Gender*]

Q: "What is your gender?"

A: "Male", "Female" or "Other"

3. [*Ethnicity*]

Q: "Which of the following best describes your ethnicity?"

A: “White”, “Black”, “Hispanic”, “Asian” or “Other”

4. [*Degree*]

Q: “What is your highest obtained degree?”

A: “None”, “High school diploma”, “Bachelor’s degree”, “Master’s degree” or “Doctorate degree”

5. [*Employment Status*]

Q: “What is your current employment status?”

A: “Full time employed”, “Part time employed”, “Unemployed”, “Self-employed”, “Student”, “Housewife / husband” or “Retired”

6. [*Income*]

Q: “Approximately, how high is your yearly income in US Dollar (after taxes)?”

A: “\$0 - \$10,000”, “\$10,000 - \$25,000”, “\$25,000 - \$50,000”, “\$50,000 - \$75,000”, “\$75,000 - \$100,000”, “\$100,000 - \$150,000” or “More than \$150,000”

7. [*Community*]

Q: “Which type of community do you live in?”

A: “Urban”, “Suburban” or “Rural”

8. [*Religion*]

Q: “How would you describe your religious views?”

A: “Christian”, “Jew”, “Buddhist”, “Muslim”, “Hindu”, “Atheist” or “Other”

9. [*Politics*]

Q: “How would you describe your political affiliation?”

A: “Very liberal”, “Liberal”, “Moderate”, “Conservative” or “Very conservative”

D.8 User Experience Questionnaire

Screen 1:

Information:

Before we inform you about your earnings and transfer you back to Prolific, we would like to ask you some additional questions. In doing so, please consider that all your data is collected anonymously and will not affect your payment!

1. [*Device*]

Q: “What device did you use during this study?”

A: “Desktop PC”, “Notebook”, “Tablet”, “Cellphone” or “Other”

2. [*Browser*]

Q: “What browser did you use during this study?”

A: “Chrome”, “Safari”, “Firefox”, “Other” or “Unknown”

The following questions are optional and can be left blank:

3. [*Understanding Problems*]

Q: “If you had any problems of understanding during the experiment, please let us know in this field:”

A: Open form text input

4. [*Technical Problems*]

Q: “If you had any technical problems during the experiment, please let us know in this field:”

A: Open form text input

5. [*Comments*]

Q: “Do you have any further comments?”

A: Open form text input

D.9 Behavioral Questionnaire

Screen 1:

Information:

Before we inform you about your earnings and transfer you back to Prolific, we would like to ask you some additional questions. In doing so, please consider that all your data is collected anonymously and will not affect your payment!

1. [*Alternative Group Identification Measure*]

Q: “On a scale from 1 to 10, please rate how closely attached you felt to Group A throughout the experiment (1: Not at all / 10: Very strongly)”

A: Any integer number between 1 and 10 (including 1 and 10)

2. [*Reason for Rule Following*]

Q: “In Part 3, why did or didn’t you follow the rule to put all the balls into the blue bucket?”

A: Open form text input

References

- Clark, C. B., Swails, J. A., Pontinen, H. M., Bowerman, S. E., Kriz, K. A., & Hendricks, P. S. (2017). A behavioral economic assessment of individualizing versus binding moral foundations. *Personality and Individual Differences, 112*, 49–54. doi:[10/gfxk2c](#)
- Cohn, A., & Maréchal, M. A. (2016). Priming in economics. *Current Opinion in Psychology, 12*, 17–21. doi:[10/gjtp9g](#)
- Dehghani, M., Johnson, K., Hoover, J., Sagi, E., Garten, J., Parmar, N. J., ... Graham, J. (2016). Purity homophily in social networks. *Journal of Experimental Psychology: General, 145*(3), 366–375. doi:[10/f8bqq8](#)
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. doi:[10/bxjdcg](#)
- Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology, 101*(2), 366–385. doi:[10/cq64hc](#)
- Jiménez, J., Gómez, Á., Buhrmester, M. D., Vázquez, A., Whitehouse, H., & Swann, W. B. (2016). The dynamic identity fusion index: A new continuous measure of identity fusion for web-based questionnaires. *Social Science Computer Review, 34*(2), 215–228. doi:[10/f8cwpq](#)
- Kimbrough, E. O., & Vostroknutov, A. (2016). Norms make preferences social. *Journal of the European Economic Association, 14*(3), 608–638. doi:[10/f8wj56](#)
- McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 105–142). New York: Academic Press.
- Swann, W. B., Gómez, Á., Seyle, D. C., Morales, J. F., & Huici, C. (2009). Identity fusion: The interplay of personal and social identities in extreme group behavior. *Journal of Personality and Social Psychology, 96*(5), 995–1011. doi:[10/dx5mf2](#)