

Responsibility Attribution Under Institutional Variation: A Robustness-Oriented Replication in Sequential Group Decisions

Janina Kraus

WISSENSCHAFTSRAUM VERHALTENSÖKONOMIK
UND GESELLSCHAFTLICHE TRANSFORMATION



Gefördert durch

zukunft.
niedersachsen

Ein Förderprogramm des Niedersächsischen
Ministeriums für Wissenschaft und Kultur
und der VolkswagenStiftung.

Responsibility Attribution Under Institutional Variation: A Robustness-Oriented Replication in Sequential Group Decisions

Janina Kraus*

Monday 15th June, 2026

Abstract

Reliable cumulative knowledge requires that behavioral findings remain stable not only across statistical replications but also under theoretically meaningful institutional variations. This paper investigates the robustness of responsibility attribution in sequential group decisions, using the experimental paradigm of Bartling et al. (2015) for the broader meta-scientific question of whether established experimental results hold under controlled design changes. We systematically extend the sequential dictator game with punishment along two institutional dimensions: (1) the presence of a default option (equal vs. unequal allocation preselected) and (2) the group-formation mechanism (random assignment vs. assignment based on university entrance grades). This design allows us to assess whether the original finding – that pivotal decision-makers receive the highest punishment – is robust to institutional variation. We cannot reproduce the dominance of pivotality as the primary driver of punishment. Instead, being the first decision-maker to choose the unequal allocation (initiation) emerges as the stronger punishment motive in our extended setting. Default structures affect punishment directed at pivotal decision-makers, but these effects disappear once punishment motives are jointly controlled. The group-formation mechanism does not significantly influence punishment behavior. These findings demonstrate that responsibility attribution is more context-sensitive than the original study suggests, and that seemingly minor procedural variations can meaningfully alter observed punishment patterns. Therefore, our study illustrates the value of robustness-oriented replication as a complement to exact replication for assessing the reliability of experimental evidence.

Keywords: responsibility attribution, replication, robustness, defaults, pivotality, group-formation mechanism

JEL classification: B41, C92, D71, D91

*Corresponding author: Janina Kraus, Leuphana University Lueneburg, Universitaetsallee 1, 21335 Lueneburg janina.kraus@leuphana.de. We thank Stefan Traub, Fabian Paetzel, Ute Losch, and the participants of the economics workshop at the Helmut Schmidt University, TU Clausthal, the GfW Conference 2021 in Magdeburg, and the 2022 ESA World Meeting in Boston.

1 Introduction

Reliable evidence is fundamental to cumulative knowledge in the social sciences. Concerns about a widespread credibility crisis have motivated sustained efforts to improve transparency, reproducibility, and methodological rigor across disciplines such as economics and psychology (Camerer et al., 2016; Maniadis et al., 2017). While practices that emphasize open data, preregistration, and robust experimental protocols have become increasingly common, an equally important question remains: whether established empirical findings are robust to theoretically meaningful yet procedurally subtle variations in experimental design. Answering this question requires not only exact replications but also robustness-oriented replications that deliberately vary key institutional features while preserving the core decision environment.

Responsibility attribution in group decision-making provides a particularly well-suited domain for this meta-scientific question. The domain is substantively important – responsibility attribution plays a central role in economic, organizational, and political contexts – yet its empirical foundations rest on a relatively small number of experimental studies. Previous experimental studies show that responsibility in group decision-making is not evenly distributed across decision-makers. In the sequential dictator game with punishment, pivotal decision-makers – those whose choices ultimately determine the outcome – are typically punished more than non-pivotal group members, indicating that pivotality is a key determinant of attributed responsibility (Bartling et al., 2015; Bhatia et al., 2024). If this pivotality finding is a robust behavioral regularity, it should persist under institutional variations that are theoretically relevant but procedurally peripheral to the core mechanism. If it does not, this reveals that the underlying attribution heuristics are more context-sensitive than assumed.

This paper contributes to this agenda by building on the sequential dictator game with punishment originally introduced by Bartling et al. (2015), situating it explicitly within the reliability and robustness research tradition. In the original design, three decision-makers sequentially choose between an equal and an unequal allocation for their group of six participants, after which recipients can assign costly punishment to each decision-maker individually. Punishment is interpreted as a behavioral expression of responsibility attribution: holding a decision-maker responsible is assumed to translate directly into the assignment of punishment points proportional to perceived responsibility (see e.g., Bartling et al., 2015; Bhatia et al., 2024). A central finding of Bartling et al. (2015) is that the pivotal decision-maker is punished most severely.

Our study adopts this experimental design but introduces two systematic extensions to assess whether this result holds under plausible and theoretically relevant variations. First, we vary the process through which participants are assigned to the roles of decision-makers and recipients, implementing group-formation mechanisms that differ in perceived legitimacy. Second, we introduce a default option that can be passively implemented unless actively overturned, allowing us to examine how active versus passive decisions affect responsibility attribution. These extensions serve a dual purpose: they speak to substantive theoretical accounts of responsibility attribution, and they function as robustness checks that test the generalizability of an experimental result.

In light of ongoing concerns about inflated or context-sensitive findings, replication and systematic extension are essential for evaluating the credibility of experimental evidence (Maniadis et al., 2017). Although definitions of replication vary, a common thread is the assessment of whether core hypotheses withstand scrutiny beyond their original implementation. While Levitt and List (2009) emphasize procedural fidelity, Berry et al. (2017) define replication more broadly as any

study that directly evaluates the validity of an original paper’s main hypothesis. Following this latter perspective, our study is designed as a robustness-oriented replication that deliberately alters key institutional features while preserving the underlying decision problem.

80 As a result, we cannot replicate the main finding of Bartling et al. (2015) in our extended setting. Rather than pivotality per se, punishment is primarily driven by the act of choosing the unequal allocation as the first decision-maker, while responsibility attribution is also systematically affected by the presence of a default option. Crucially, however, our results should not be interpreted as a failed replication in the narrow sense. Rather, they illustrate precisely why robustness-oriented
85 replications are indispensable for cumulative scientific progress: the original findings remain informative and partially supported, yet our extensions demonstrate that seemingly (minor) procedural variations can meaningfully alter observed punishment patterns. These findings underscore the importance of robustness checks and show that responsibility attribution in sequential group decisions is more context-dependent than previously assumed.

90 The remainder of the paper is organized as follows. Section 2 reviews the related literature. Section 3 describes the experimental design, procedures, and hypotheses. Section 4 presents the results and Section 5 concludes.

2 Further Literature

Beyond its methodological contribution to the replication debate, this study connects to three substantive strands of literature: responsibility attribution in collective decision-making, the behavioral effects of defaults and acts of omission, and the role of legitimate status differences. Importantly, these strands are not independent themes of equal weight in this paper. Rather, they collectively motivate the institutional variations this study introduces to assess whether an established experimental finding holds beyond its original design. If responsibility attribution
100 depends systematically on institutional features such as default options or the perceived legitimacy of decision-makers, then robustness cannot be assessed solely by replicating outcomes; it must also account for the stability of the underlying mechanisms.

Replication and Reliability in Experimental Research.

105 In recent years, replications have attracted increasing attention in the social sciences, prompting systematic efforts to assess both the frequency and the success of replication attempts. Berry et al. (2017) examine all empirical articles published in the 2010 volume of the American Economic Review and document that while a majority of papers receive follow-up studies in the form of robustness checks or extensions, explicit replication attempts remain relatively rare. Only 29 percent of the original studies are subject to replication, with highly cited papers being more likely to attract such efforts.

Complementary evidence is provided by Maniadis et al. (2017), who analyze replication practices across the top 150 economics journals and report that only about 4.2 percent of experimental studies are replicated. Notably, however, nearly half of these replication attempts successfully
115 reproduce the original findings, suggesting that replications – when conducted – can be informative despite their low prevalence.

Evidence from psychology paints a more cautious picture. The Reproducibility Project in Psychology replicates 100 studies from leading journals and finds substantially weaker support for

original results across multiple criteria, including statistical significance and effect sizes (Collabo-
120 ration, 2015). In contrast, Camerer et al. (2016) report comparatively high replication success rates
for laboratory experiments in economics, a pattern they attribute to the high level of transparency
typical for experimental designs, including the availability of data and instructions.
These findings suggest that replication success depends not only on reproducing statistical results,
but also on the stability of the behavioral mechanisms under investigation. This highlights the im-
125 portance of robustness checks that go beyond outcome replication and explicitly test whether core
attribution mechanisms persist under changes in contextual features, as this paper does.

Responsibility Attribution in Collective Decision-Making.

Responsibility attribution in group decision-making is a theoretically well-developed domain (see
130 e.g., Chockler and Halpern, 2004; Spellman, 1997; Engl, 2018), making it an ideal topic for
robustness-oriented replication. The experimental literature has identified multiple heuristics that
recipients use when assigning punishment – such as pivotality and agenda power – and the coexis-
tence of these heuristics raises natural questions about which mechanisms dominate under which
institutional conditions.

135 Recipients of group decisions typically do not distribute punishment evenly across decision-makers,
but condition it on individuals’ perceived influence on the outcome. In the sequential dictator game
introduced by Bartling et al. (2015), pivotal decision-makers – those whose choice ultimately de-
termines the unequal outcome – receive the highest punishment, suggesting that pivotality serves
as a salient heuristic for responsibility attribution.

140 Subsequent studies extend this framework and highlight additional dimensions that shape punish-
ment behavior. Bhatia et al. (2024) replicate the pivotality result in larger groups and show that
punishment strategies are heterogeneous, with some participants focusing on pivotality, others on
the choice of unequal allocations, and a substantial fraction assigning little or no punishment. Ca-
puto and Duch (2024) extend the experimental design of Bartling et al. (2015) with a proposer
145 stage and find that fairness concerns about the proposed allocation drive punishment of proposers
more strongly than considerations of pivotality. Other work emphasizes the role of agenda power.
Duch et al. (2014) demonstrate that punishment is concentrated on the decision-maker who holds
both proposal power and the greatest voting weight, while neither veto power nor proportional
vote share reliably drives targeting behavior. Similar patterns emerge in committee settings, where
150 agenda control increases acceptance rates and is not punished per se, but evaluated relative to
individual benefits (Fischbacher and Schudy, 2020).

Beyond institutional features, responsibility attribution has also been shown to depend on the iden-
tity and group membership of decision-makers. For instance, Grimm and Klimm (2019) demon-
strate that the group affiliation of a decision-maker — specifically whether they belong to an in- or
155 outgroup — systematically shapes how responsibility is assigned, even when controlling for actual
performance differences.

From a robustness perspective, the literature indicates that responsibility attribution is driven not
by pivotality alone, but by a plurality of potentially competing heuristics. The coexistence of these
heuristics raises concerns about the generalizability of punishment patterns across institutional set-
160 tings, and underscores the need for robustness-oriented extensions in this domain.

Defaults, Acts of Omission, and Responsibility.

A key institutional feature that has received comparatively little attention in group settings is the

165 presence of default options. Conceptually, a default differs from the status quo in that it reflects the framing or presentation of alternatives rather than an outcome that has already been implemented (Samuelson and Zeckhauser, 1988; Johnson et al., 1993). Both defaults and status quo options can be realized actively, through acts of commission, or passively, through acts of omission. Acts of omission – defined as instances in which an expected action is not taken – are often preferred to acts of commission, particularly when decisions involve potentially harmful consequences (Shultz et al., 1981; Spranca et al., 1991). Importantly, acts of omission and default-based outcomes are not only evaluated differently in choice behavior, but also in moral judgment and punishment. Omissions are typically perceived as less blameworthy and are punished less severely than comparable acts of commission (Spranca et al., 1991; DeScioli et al., 2011a). This asymmetry suggests that default options may systematically attenuate responsibility attribution by obscuring causal agency and underlying intentions.

175 This omission bias is commonly explained by the tendency to attribute responsibility more strongly to active choices than to inaction, as intentions are more salient in cases of commission than omission (Ritov and Baron, 1992; Gärtner et al., 2017). Related literature documents robust default effects, showing that individuals are substantially more likely to accept an option when it is designated as the default, even when this option does not reflect their underlying preferences (see e.g., Johnson and Goldstein, 2003; Sunstein and Thaler, 2003; Jachimowicz et al., 2019). Closely related, individuals also display a pronounced status quo bias, favoring the preservation of existing arrangements over change (see e.g., Samuelson and Zeckhauser, 1988; Johnson et al., 1993; Gärtner, 2018).

185 While these effects are well established in individual decision-making, it remains largely unexplored whether defaults and acts of omission also influence responsibility attribution and punishment in group contexts. If passive implementation through defaults systematically attenuates perceived individual responsibility, this has direct implications for the robustness of punishment patterns observed in earlier studies.

190 *Status Differences and Legitimacy.*

Finally, responsibility attribution in group decision-making is closely linked to status differences and their perceived legitimacy. In laboratory experiments, roles and decision power are often assigned randomly, implying that advantageous positions are obtained by chance and may be perceived as unearned or illegitimate. Experimental evidence suggests that such assignments affect both the willingness of high-status individuals to exercise authority and the acceptance of their decisions by lower-status group members (Feltovich, 2019).

Status hierarchies are socially recognized rankings that confer access to resources and preferential treatment (Ball et al., 2001). A unifying view characterizes status by relative rank, social recognition, and entitlement to resources (Henrich and Gil-White, 2001). Economic research increasingly acknowledges that status considerations shape allocation decisions and the acceptance of inequality, with higher-status individuals often perceived as more deserving and income disparities being more readily tolerated when status differences exist (Fershtman et al., 2012). Consistent with this view, high-status individuals are often treated more favorably and punished less, while low-status individuals behave more generously toward higher-status counterparts (Ball and Eckel, 2005). Experimental evidence further shows that low-status individuals are more willing to accept disadvantageous payoff distributions, indicating a greater tolerance for inequality when interacting with higher-status decision-makers (Albrecht et al., 2013).

Crucially, however, such acceptance depends on the perceived legitimacy of the status structure. When hierarchies are regarded as legitimate, unequal outcomes are more likely to be accepted by both advantaged and disadvantaged individuals; when legitimacy is lacking, status differences may provoke resistance and punishment (Tyler, 2006). The importance of legitimacy is further underscored by evidence from ultimatum games with earned versus randomly assigned roles. When status is earned through performance, proposers offer less and recipients accept lower offers, particularly for allocations below an equal split (Fleiß, 2015). Similar patterns emerge in settings with explicit status differences, where both one’s own status and the status of the interaction partner systematically affect acceptance behavior (Blue et al., 2016). These findings suggest that perceived legitimacy moderates punishment and rejection responses rather than eliminating inequality concerns altogether. From a robustness perspective, these results imply that punishment and acceptance of decisions are contingent on the perceived legitimacy of status hierarchies — calling into question whether findings from experiments with randomly assigned roles generalize to environments in which status differences are institutionally justified.

Despite the relevance of defaults, omission bias, and legitimate status differences for responsibility attribution, these factors have largely been studied in isolation. To the best of our knowledge, no experimental study jointly examines how default options and the legitimacy of status differences interact in shaping responsibility attribution in collective decision-making. Addressing this gap, our study systematically varies both dimensions to assess the robustness of responsibility attribution under institutional change.

3 Experiment

3.1 Experimental Design

In this study we use the sequential dictator game with punishment introduced by Bartling et al. (2015). In contrast to the original study, our experimental design systematically combines two default conditions (equal vs. unequal allocation preselected) with two group-formation mechanisms (random assignment vs. assignment based on university entrance grades), resulting in a 2x2 between-subjects design (see Table 1).

Table 1: 2×2 Between-Subjects Experimental Design

Group-formation mechanism	Default condition	
	Equal	Unequal
Random	T1: Random-equal ($n = 72$)	T2: Random-unequal ($n = 66$)
University entrance grade	T3: Grade-equal ($n = 66$)	T4: Grade-unequal ($n = 66$)

Notes: Total $N = 270$ participants, yielding 2,430 individual punishment decisions (3 recipients \times 6 rounds \times 135 decisions).

The experiment consisted of four parts. Instructions for Parts 1 and 2 were distributed and read aloud together, as were the instructions for Parts 3 and 4. Throughout the experiment, neutral language was used deliberately, avoiding potentially loaded terms such as punishment or status. The design is illustrated schematically in Figure 1. Instructions are presented in the Online Appendix B.

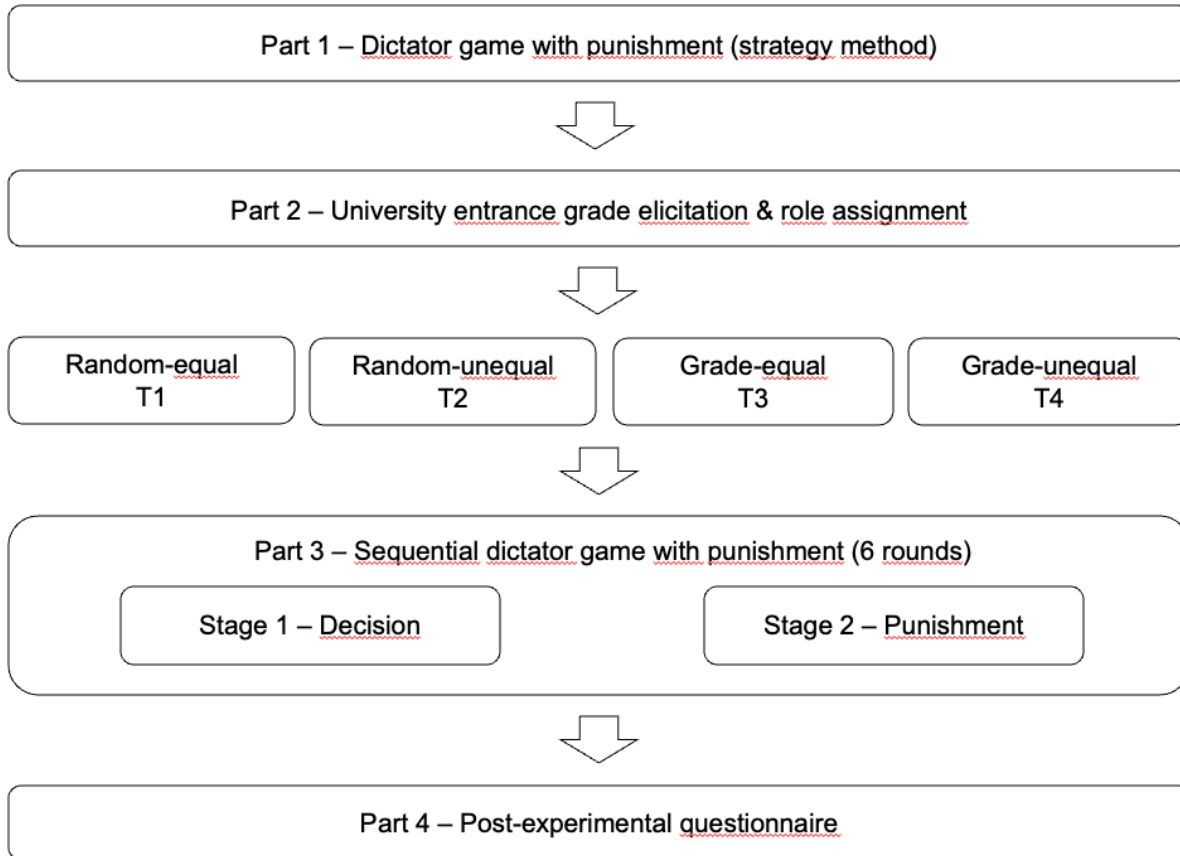


Figure 1: Stages of the Experiment

At the beginning of the experiment (Part 1), participants took part in a standard dictator game with punishment using the strategy method. Each participant acted once as dictator and once as recipient. Dictators chose between two possible distributions, (5,5) and (9,1). Recipients could assign up to five punishment points to the dictator, where each punishment point reduced the dictator's payoff by one point and imposed a cost of one point on the recipient. At the end of the experiment, one role was randomly chosen and the corresponding decisions were implemented.

In Part 2, participants reported their university entrance grade or an equivalent qualification. Importantly, this information was collected in all treatments, including those with random role assignment, to ensure maximal comparability across conditions. Between Parts 2 and 3, participants were assigned their fixed roles for the remainder of the experiment. Decision-makers were assigned to role A and recipients to role B. In the random treatments, this assignment was purely random. In the grade-based treatments, participants with a university entrance grade above the session median were assigned to role A, while those below the median were assigned to role B. The

255 assignment procedure was common knowledge to all participants after providing their university
entrance grade (see Table 8 in the Online Appendix A.1 for mean university entrance grades in
each treatment).

The selection of these two mechanisms – random assignment and assignment based on university
entrance grades – is based on the assessments of other participants in a pre-study (Kraus and Traub,
260 2026). In the pre-study, other participants were asked to evaluate four different mechanisms in
terms of their legitimacy, which then determine the roles (decision-maker and recipient) in the
decision-making situation used here. This study then utilized the most legitimate (random) and the
least legitimate (university entrance grades) mechanisms.

Part 3 constituted the core of the experiment and is adapted from Bartling et al. (2015). At its
265 outset, participants were informed about their roles and about the two-stage structure of the up-
coming decision task. In each round, participants were organized into groups of six, consisting of
three decision-makers and three recipients. While group composition changed after each round,
individual roles remained fixed. At the beginning of each round, decision-makers were informed
about their position within the sequence: the first decision-maker (A1), the second decision-maker
270 (A2), and the third decision-maker (A3). The two stages described below were repeated for a total
of six rounds.

In the first stage, the three decision-makers determined how to allocate 30 points among the six
group members. Each decision-maker sequentially chose between two distributions: an equal allo-
cation (5,5,5,5,5,5) and an unequal allocation (9,9,9,1,1,1). Depending on the treatment condition,
275 one of these distributions was preselected as the default option. Each decision-maker had 30 sec-
onds to make a choice. If no active choice was made within this time, the preselected distribution
was recorded. As soon as at least two decision-makers selected the same distribution, it was im-
plemented for the entire group. Decision-makers could observe all previous choices within their
group. While decision-makers selected allocations, recipients were asked to state their expectations
280 regarding the decisions of each position.

In the second stage, recipients assigned punishment points to the three decision-makers. They were
informed about (i) the implemented distribution, (ii) the default distribution, and (iii) the individual
choices of each decision-maker. Recipients could allocate up to seven punishment points across
the three decision-makers, with the restriction that no decision-maker could lose more points than
285 they had received in the allocation stage. Assigning at least one punishment point imposed a cost
of one point on the recipient. The allocation of punishment points constitutes the main measure
of responsibility attribution at the individual level. At the same time, decision-makers were asked
to estimate how many punishment points they expected to receive. For payoff determination in
Part 3, one round was randomly selected. The allocation implemented in that round was applied,
290 and one recipient was randomly chosen whose punishment allocation was executed. Participants
were informed about the selected round, the selected recipient, and the resulting payoffs only after
completion of Part 3.

Finally, Part 4 consisted of a post-experimental questionnaire. Participants provided further infor-
mation, including age, gender, field of study, and semester. They also reported their risk attitudes
295 and were asked to explain the reasoning behind their decisions in Part 3. In addition, participants
evaluated the group-formation mechanism (random vs. university entrance grade) with respect to
its perceived legitimacy and the perceived deservability of roles A and B.

Experimental studies typically elicit behavior either via the strategy method or the direct response

300 method. While Bartling et al. (2015) employ the strategy method to elicit punishment decisions, our study uses the direct response method, reflecting an intentional variation that speaks to the reliability of responsibility attribution across elicitation procedures. Existing evidence suggests that in simple decision environments such as dictator and ultimatum games, both methods yield comparable treatment effects, even if the intensity of responses may differ (Brandts and Charness, 2000).
305 Consequently, observing similar qualitative patterns across elicitation methods can be interpreted as further support for the robustness and generalizability of the underlying behavioral mechanisms.

3.2 Procedures

From September to November of 2020, the experimental sessions were carried out at the University of Hamburg. Given that each group contained precisely six individuals, the number of participants
310 in each session was always a multiple of six. For each of the four experimental treatments, three separate sessions were organized, with each session involving either 18, 24, or 30 participants. was determined by two factors: having at least 30 decision-maker and recipients in each treatment and time constraints as well as limitations on the availability of participants due to the COVID-19 lockdowns. In Table 8 in the Online Appendix A.1, we provide information on the composition
315 of the sample within each treatment group. To prevent the repetition of participation, individuals were permitted to partake in the study only once. Furthermore, any individuals who had previously participated in the pre-study (Kraus and Traub (2026)) were excluded from the laboratory sessions. Questions were encouraged and addressed after the instructions were read aloud by the experimenter.

320 The study encompassed a total of 270 participants. The majority of the subjects were undergraduate students from a variety of academic backgrounds, and they were recruited through hroot Bock et al. (2014). This yielded a total of 2,430 individual punishment decisions, derived from three recipients per group across six rounds. The experiment was programmed and conducted using oTree Chen et al. (2016), and each session lasted approximately one hour. The mean earnings of the
325 participants amounted to €14.83, which included a fixed participation fee of €5. In line with the procedure used by Bartling et al. (2015), participants were paid in experimental points, which were later converted into Euros at a 1:1 rate. Final earnings were disclosed only after the conclusion of the experiment.

3.3 Hypotheses

330 This section proceeds in two steps. The replication hypotheses (H1-H4) assess whether the core mechanisms identified by Bartling et al. (2015) are robust to our extended design. The extension hypotheses (H5-H6) test whether institutional features – defaults and group-formation mechanisms – systematically moderate responsibility attribution. By structuring the hypotheses in this way, we explicitly distinguish between replication-oriented tests of established findings and novel predic-
335 tions that examine boundary conditions of the original result. At the time this experiment was conducted, the pre-registration of hypotheses was not yet standard practice. Accordingly, the hypotheses presented below were not pre-registered.

3.3.1 Replication Hypotheses: Punishment Motives

From a purely self-interested perspective, recipients should never punish, as punishment is costly and yields no material benefit. Although extensive experimental evidence shows systematic deviations from this prediction (e.g., Fehr and Gintis, 2007; Akbaş et al., 2019), some participants may nonetheless behave in a strictly payoff-maximizing manner.

A first motive is outcome-based punishment. Inequality-averse recipients focus exclusively on the implemented allocation and assign responsibility accordingly, irrespective of individual choices. Thus, unequal outcomes should trigger punishment, whereas equal outcomes should not:

Hypothesis 1 (Outcome) *When the unequal allocation is implemented, decision-makers are punished, while no punishment is assigned for an equal outcome.*

Beyond outcomes, punishment may also be driven by intentions. Choosing the unequal allocation increases the likelihood that an unequal outcome will be implemented and is therefore perceived as unkind, even if the final outcome turns out to be equal. Consequently, decision-makers who choose the unequal allocation are expected to receive more punishment than those who choose the equal allocation:

Hypothesis 2 (Choice) *Choosing the unequal allocation leads to more punishment.*

Punishment should be particularly pronounced when a decision-maker's choice is consequential for the outcome initiating the decision sequence (Spellman, 1997). In our design, this applies especially to first decision-maker whose choices shape subsequent decisions:

Hypothesis 3 (Initiation) *More punishment points are awarded to the first decision-maker choosing the unequal allocation.*

A closely related and central mechanism in the original study is pivotality. The pivotal decision-maker – whose choice ultimately determines the implemented allocation – is expected to receive more punishment than non-pivotal decision-makers. This hypothesis directly replicates the main finding of Bartling et al. (2015) and its extensions (Bhatia et al., 2024), and is also consistent with formal responsibility models (Chockler and Halpern, 2004; Engl, 2018):

Hypothesis 4 (Pivotality) *The pivotal decision-maker receives more punishment points than the non-pivotal decision-maker.*

3.3.2 Extension Hypotheses: Defaults and Status Differences

Beyond these replication hypotheses, we examine whether institutional features – defaults and status differences – systematically moderate punishment and the strength of the mechanisms outlined above. These hypotheses test the boundary conditions of the pivotality finding and thus constitute the robustness-oriented component of our study.

A robust finding in moral psychology and experimental economics is that acts of omission are punished less than acts of commission (e.g., Spranca et al., 1991; DeScioli et al., 2011b). Since defaults can be implemented passively, they obscure intentions and reduce perceived responsibility. Accordingly, choosing an unequal allocation should be punished less when it coincides with the default:

Hypothesis 5 (Default) *Choosing the default leads to less punishment.*

Status differences constitute a second extension. High-status decision-makers are typically perceived as more competent and deserving, and inequalities generated by legitimate procedures are more readily accepted (Albrecht et al., 2013; Fleiß, 2015). When status differences are seen as
380 legitimate, unequal outcomes should therefore trigger less punishment:

Hypothesis 6 (Status) *A more legitimate status structure leads to higher acceptance of inequalities, so unequal choices of randomly assigned decision-makers are less punished.*

Finally, the interaction between defaults and status differences yields ambiguous predictions, as no prior study combines both dimensions in a punishment context. We conjecture that punishment
385 will be lowest when an unequal allocation can be implemented passively by high-status decision-makers in a random (and thus more legitimate) group-formation mechanism, and highest when unequal allocations must be actively chosen under less legitimate status structures. However, whether defaults or status differences exert the stronger influence on punishment remains an open empirical question.

390 **4 Results**

The structure of this section follows a two-step approach aligned with our hypothesis structure. We begin by closely aligning our analysis with that of Bartling et al. (2015), thereby conducting a replication in the narrow sense and testing Hypothesis 1 to 4 (Section 4.1). We then turn to the extensions introduced in the present study – the treatment variations involving defaults and
395 group-formation mechanisms – testing Hypothesis 5 and 6 (Section 4.2).

4.1 Comparison to Bartling et al. (2015)

Since the experimental design is adapted from Bartling et al. (2015), this section compares the results of our study with those reported therein. First, we discuss the punishment points assigned to each individual decision-maker. Second, we present an econometric comparison of the different
400 punishment motives. Although our primary focus is not on the behavior of decision-makers, we also provide an overview of their choices and contrast them with the findings of Bartling et al. (2015). To compare the overall effects of the different punishment motives, this section draws on the aggregated data of the present study. The analysis uses the publicly available dataset and corresponding do-file from Bartling et al. (2015) – no changes are made to the dataset.

405 **4.1.1 Responsibility Attribution for Decision-Makers**

Table 2 reports the average punishment points assigned to each decision-maker across all eight possible decision sequences in both studies. Owing to methodological differences – specifically, the use of the direct response method in our study versus the strategy method employed by Bartling et al. (2015) – the underlying data structures differ, necessitating different statistical tests. While
410 Bartling et al. (2015) employ Wilcoxon signed-rank tests, we use Wilcoxon rank-sum tests due to

the independence of observations in our setting. Additional analyses based on restricted datasets are reported in Table 9 and Figure 3 in the Online Appendix A.2.

Table 2: Comparison to Bartling et al. (2015)

Allocation	Decision sequence	Decision-maker 1	Decision-maker 2	Decision-maker 3	Pivotal vs. all	Pivotal vs. initiator	Initiator vs. all
		Bartling et al. (2015)			Wilcoxon sign rank		
Unequal	u-u-u	1.50	1.85	0.86	$p < 0.001$	$p = 0.294$	
	u-u-e	1.86	1.92	0.26	$p < 0.001$	$p = 0.960$	
	u-e-u	1.68	0.07	2.39	$p < 0.001$	$p = 0.006$	
	e-u-u	0.11	1.83	2.33	$p < 0.001$	$p = 0.012$	
Equal	u-e-e	1.33	0.10	0.08			
	e-u-e	0.17	1.43	0.08			
	e-e-u	0.06	0.03	0.92			
	e-e-e	0.08	0.07	0.03			
		This study			Wilcoxon rank-sum		
Unequal	u-u-u	1.33	1.00	0.87	$p = 0.946$	$p = 0.765$	$p = 0.628$
	u-u-e	2.00	2.02	0.09	$p = 0.002$	$p = 0.782$	$p < 0.001$
	u-e-u	2.67	0.13	2.01	$p = 0.025$	$p = 0.201$	$p < 0.001$
	e-u-u	0.00	2.64	1.58	$p = 0.129$	$p = 0.009$	$p < 0.001$
Equal	u-e-e	1.68	0.09	0.12			
	e-u-e	0.05	1.41	0.05			
	e-e-u	0.03	0.08	2.26			
	e-e-e	0.18	0.13	0.17			

Table notes: “u” denotes a choice of the unequal allocation; “e” denotes a choice of the equal allocation. The three columns on the right report p-values from Wilcoxon signed-rank tests for Bartling et al. (2015) and from rank-sum tests for the present study. These tests compare punishment assigned to the pivotal decision-maker with punishment assigned either to the two remaining decision-makers (“pivotal vs. all”) or solely to the initiator (“pivotal vs. initiator”). In addition, the present study includes a comparison of punishment assigned to the initiator with punishment assigned to the two remaining decision-makers (“initiator vs. all”).

As expected, punishment is generally lower following equal choices than unequal ones, irrespective of the resulting outcome. In the present study, equal choices receive an average punishment of 0.14 points, whereas unequal choices are punished with 1.89 points on average, a difference that is highly significant (rank-sum test, $p < 0.001$). Thus, the results provide support for Hypothesis 2. In contrast to Hypothesis 1, however, punishment points are assigned even in cases where the equal outcome is implemented. Although punishment is directed primarily toward the decision-maker who had chosen the unequal allocation, a non-negligible number of punishment points is also observed in the e-e-e sequence (see Table 2). This divergence from purely outcome-based punishment suggests that intention-related cues play a robust role in responsibility attribution. Consequently, Hypothesis 1 must be rejected.

Result 1 (Outcome and Choice) *Equal choices are punished significantly less frequently than unequal choices. Nevertheless, opting for the equal allocation does not fully shield recipients from punishment. A non-zero share of punishment points is observed even in situations involving equal choices and equal outcomes in both studies.*

Given that unequal outcomes and the associated punishment are of primary interest, we now focus on the upper panel of Table 2, where unequal allocations arise.

Sequences 1 (u-u-u) and 2 (u-u-e), in which both decision-makers initially select the unequal allocation, display a high degree of consistency across the two studies. In both settings, first and second decision-makers receive similar levels of punishment, with no statistically significant differences between them. Accordingly, we find no evidence that either the first (i.e., initiator) or the second (i.e., pivotal) decision-maker is punished more heavily in these sequences.

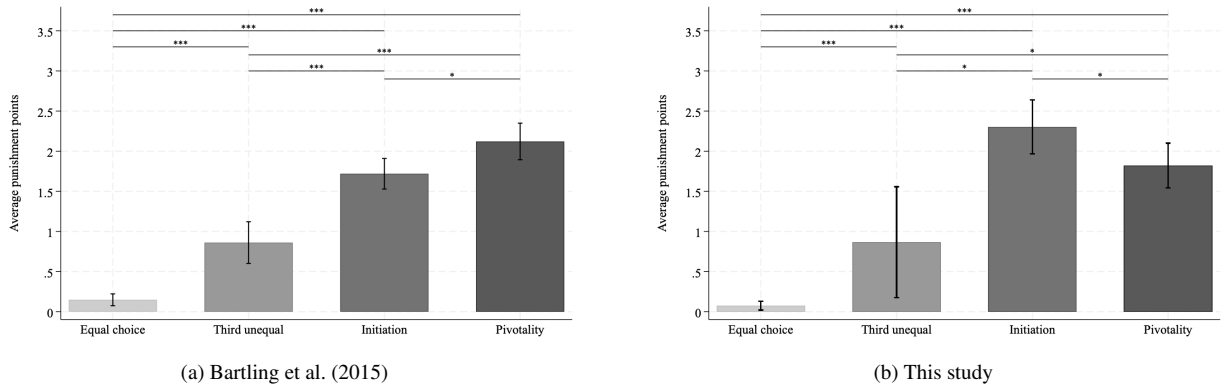
In contrast, Sequences 3 (u-e-u) and 4 (e-u-u) reveal substantial differences between the studies. The central result reported by Bartling et al. (2015) – that the pivotal decision-maker is punished significantly more than the other decision-maker – is not replicated in our data. Instead, the first decision-maker who selects the unequal allocation receives significantly more punishment than the pivotal decision-maker.

This pattern suggests that pivotality and initiation appear to exert a similar influence in some decision contexts, whereas in others their effects diverge substantially.

To further investigate punishment behavior across both studies, Figure 2 reports average punishment points for unequal outcomes, distinguishing between four categories of decision-makers: equal choice, third unequal choice, initiator, and pivotal decision-maker. Additional details on the classification of decision-makers are provided in Table 10 in the Online Appendix A.2.

Although Bartling et al. (2015) label the initiator as “unkind” in their figure, the underlying definition is equivalent: a decision-maker who selects the unequal allocation and can still influence the outcome without being pivotal. To maintain a consistent linguistic distinction between an unkind intention — which includes both initiator and pivotality (see Table 10 in the Online Appendix A.2) — we use the term initiator for the first decision-maker in a sequence choosing unequally.

As outlined above, the central difference between the two experiments lies in the punishment of the first and the pivotal decision-maker. As shown in Figure 2, Bartling et al. (2015) find that the pivotal decision-maker receives significantly more punishment points than the first decision-maker (i.e., the initiator). In contrast, our results reverse this pattern: the initiator receives the highest level of punishment, with a mean of 2.30 points, while the pivotal decision-maker receives the second most punishment (1.82). Punishment following equal choices is minimal and consistent across both studies. The same holds for the third decision-maker who selects the unequal allocation after both preceding decision-makers have already done so.



Notes: The figure displays the mean number of assigned punishment points (ranging from 0 to 7) across four categories of decision-makers. *Equal choice* refers to choosing the equal allocation when the two other decision-makers choose the unequal allocation. *Third unequal* denotes the third decision-maker choosing the unequal allocation after the two preceding decision-makers have also chosen unequal. *Initiation* refers to the first decision-maker in a sequence choosing the unequal allocation, whereas *Pivotality* denotes the pivotal decision-maker ultimately determining the unequal outcome. Bars indicate 95% confidence intervals. The p-values reported in the left panel are based on signed-rank tests, while those in the right panel are based on rank-sum tests. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Figure 2: Average Punishment for an Unequal Outcome

Figure 4 in the Online Appendix A.2 further illustrates the distribution of punishment points for initiators and pivotal decision-makers. In both studies, zero punishment occurs at similar rates for both roles. However, conditional on punishment, initiators in the present study are substantially more likely to receive high punishment levels, including the maximum punishment of seven points. Taken together, the descriptive comparison indicates that pivotality remains important, but no longer dominates responsibility attribution as strongly as in the original study. Instead, punishment behavior reflects a more balanced weighting between initiation and pivotality. From a robustness perspective, these findings illustrate that the dominance of pivotality as the primary attribution heuristic — the central result of Bartling et al. (2015) — does not replicate under our extended design. Rather than a uniform shift in punishment patterns, the divergence is concentrated in specific decision sequences, suggesting that pivotality and initiation function as competing heuristics whose relative salience depends on the structure of the decision context.

4.1.2 Econometric Comparison of Different Punishment Motives

In this section, we jointly examine the effects of different punishment motives. To ensure comparability, we adopt the same regression model and punishment variables as Bartling et al. (2015). As noted, Bartling et al. (2015) do not explicitly include the initiator in their econometric specification. Instead, they define unkind intention to encompass all decision-makers who choose the unequal allocation while the final outcome remains undetermined – a definition that includes both pivotal decision-makers and initiators. Building on the previous section, we incorporate this differentiation directly into the empirical analysis to achieve a clearer separation of underlying motives. In Models 1 through 6, each punishment motive is included separately as a predictor of assigned punishment points. Model 7 includes all motives simultaneously. The final column of Table 3 reports the estimates from Bartling et al. (2015) for direct comparison. All OLS regressions use

cluster-robust standard errors at the individual level. To assess robustness, we estimate additional specifications, including regressions restricted to the first round and to the random treatments (see Table 11 in the Online Appendix A.2).

Table 3: Comparison of Different Punishment Motives

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	Bartling et al. (2015)
Outcome unequal	1.134*** (0.133)						-0.068 (0.053)	0.048 (0.070)
Unequal choice		1.756*** (0.163)					0.792* (0.325)	0.782*** (0.137)
Intention unkind			1.711*** (0.161)					
Initiator				1.697*** (0.172)			0.748 (0.397)	0.517** (0.196)
Outcome unequal X Initiator					1.913*** (0.231)		0.688* (0.299)	0.289 (0.232)
Pivotal						1.397*** (0.191)	0.954** (0.326)	1.209*** (0.203)
Constant	0.290*** (0.049)	0.138*** (0.041)	0.185*** (0.043)	0.274*** (0.043)	0.389*** (0.048)	0.424*** (0.048)	0.143** (0.044)	0.083* (0.037)
R^2	0.117	0.297	0.260	0.201	0.130	0.069	0.310	0.281
N	2430	2430	2430	2430	2430	2430	2430	1728

Table notes: Random effects OLS regression with punishment points for decision-makers as dependent variable (from 0 to 7). All punishment variables are dummy variables. *Outcome unequal* equals 1 if the unequal allocation is implemented. *Unequal choice* equals 1 if the unequal allocation is chosen. *Intention unkind* equals 1 if a decision-maker opts for the unequal allocation without having a majority. *Initiation* equals 1 if a decision-maker is the first one in a decision sequence choosing the unequal outcome. *Pivotal* equals 1 if a decision-maker is pivotal for the unequal outcome. Standard errors are in parentheses and clustered on 270 individuals in this study and 72 individuals in Bartling et al. (2015). *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

485 Several conclusions emerge from Table 3. First, each punishment motive, when considered in iso-
 lation, has a statistically significant positive effect, increasing assigned punishment by at least one
 point. The largest effect is associated with being the initiator of an unequal outcome (coefficient
 1.913), implying that initiators receive, on average, 1.91 additional punishment points relative to
 non-initiators. Second, unequal choice exhibits the greatest explanatory power, followed by unkind
 490 intention. Third, initiation displays substantially greater explanatory power than pivotality (R^2 of
 0.201 and 0.130 versus 0.069, respectively).

Model 7 includes all punishment motives simultaneously. In both studies (Model 7 and 8), unequal
 choice remains statistically significant. While both initiation and pivotality significantly affect
 punishment, in our data the effect of being the initiator for the unequal outcome is only weakly
 495 significant at the 5% level. By contrast, being pivotal remains (highly) significant in both studies.
 The regression analysis shows that all punishment motives significantly increase punishment when
 considered individually, with initiation having the strongest effect. However, once all motives

are included jointly, unequal choice and pivotality remain robust predictors, while the effect of initiation weakens. Compared to Bartling et al. (2015), the results confirm the importance of pivotality but highlight that distinguishing initiation as a separate motive refines the interpretation of punishment behavior. Thus, both Hypothesis 3 and Hypothesis 4 receive empirical support.

Result 2 (Initiation and Pivotality) *Both punishment motives significantly affect the allocation of punishment points. Depending on the decision context, the two motives are either equally relevant or punishment is directed somewhat more strongly toward the initiator.*

Importantly, the fact that initiation emerges as a distinct and influential motive – not captured in the original specification – illustrates how seemingly minor differences in empirical modeling choices can affect which behavioral regularities become visible, and underscores the value of extended replication designs.

4.1.3 Decision-Maker's Choices

Although our primary focus is not on the behavior of decision-makers, it is informative to briefly summarize their choices and compare them to those reported by Bartling et al. (2015). In this study, 135 participants are decision-maker and are randomly assigned each round to one position within the decision sequence. Of the 270 group observations (45 groups for 6 rounds), 54 (20.0%) lead to unequal outcomes, while at the individual level, unequal allocations are chosen in 175 out of 810 decisions (21.6%). These rates stand in stark contrast to the findings of Bartling et al. (2015), where nearly two-thirds of all decisions result in unequal outcomes.

Table 4 summarizes decision-makers' choices along with the corresponding expected punishment points at each decision node. Columns (3) and (4) present the results from this study, whereas Columns (5) and (6) report the corresponding results from Bartling et al. (2015).

Relative to Bartling et al. (2015), decision-makers in the present study behave substantially more egalitarian. Only 20.0% of group decisions result in unequal outcomes, compared to 67.4% in the original study. Similarly, unequal allocations are chosen in only 21.6% of all individual decisions. These differences emerge throughout the decision process. First decision-makers choose unequal allocations less frequently, second decision-makers defer unequal outcomes more often, and pivotal (third) decision-makers are less likely to implement unequal allocations than in Bartling et al. (2015).

The underlying drivers of this divergence cannot be identified conclusively. One possible explanation lies in methodological differences: Bartling et al. (2015) employ the strategy method, while the present study relies on the direct response method. Alternatively, the discrepancy may stem from the implementation of defaults in the present design, or from broader shifts in behavioral norms. Such dynamics could emerge if different choices are perceived as normatively appropriate across experimental contexts. Regardless of the precise mechanism, these findings once again underscore the importance of replication and systematic extensions in experimental research – and signal that punishment norms may themselves be endogenous to the decision environment.

Table 4: Choices and Expected Punishment at each Decision-Node

Decision	Sequence (decision in bold)	This study		Bartling et al. (2015)	
		Choices	Exp. punishment	Choices	Exp. punishment
1	u ...	22.59%	1.99	58.3%	1.70
	e ...	77.41%	0.14	41.7%	0.11
2	u-u ...	39.3%	1.81	59.7%	1.90
	u-e ...	60.7%	0.11	40.3%	0.08
	e-u ...	20.1%	1.85	61.1%	1.68
	e-e ...	79.9%	0.13	38.9%	0.07
3	u-u-u	20.8%	0.87	22.2%	0.86
	u-u-e	79.2%	0.09	77.8%	0.26
	u-e-u	40.5%	2.09	70.8%	2.39
	u-e-e	59.5%	0.12	29.2%	0.08
	e-u-u	35.7%	1.58	62.5%	2.33
	e-u-e	64.3%	0.49	37.5%	0.08
	e-e-u	7.8%	2.26	1.40%	0.92
	e-e-e	92.2%	0.17	98.6%	0.03
Unequal outcome		20.0%		67.4%	

Table notes: “u” denotes a choice of the unequal allocation; “e” stands for a choice of the equal allocation.

535 4.2 Extension of Bartling et al. (2015): Treatment Variations

Building on the aggregated results presented above, we now turn to treatment-level differences and examine whether the default, the group-formation mechanism, or the interaction of both systematically affect punishment behavior – thereby testing Hypothesis 5 and Hypothesis 6. This analysis constitutes the robustness-oriented core of our study: only by varying institutional features can we assess whether the punishment patterns documented in Section 4.1 reflect stable behavioral regularities or context-sensitive responses. We proceed in two steps: first, we compare behavior across the four treatments; second, we analyze the presence of a default effect and status differences.

4.2.1 Punishment Behavior across Treatments

Table 5 provides an overview of the punishment points assigned in this study, with a focus on initiators and pivotal decision-makers across treatment conditions.

Focusing first on overall punishment levels, we find that they are largely comparable across conditions, both between defaults and group-formation mechanisms, as well as across the four treatment variations. The grade-unequal treatment exhibits slightly lower average punishment points than the other treatments; however, these differences are not statistically significant. In this section, we present the results of the t-test, with rank-sum tests yielding comparable levels of significance.

Turning to differences in punishment assigned to the initiator and the pivotal decision-maker, the aggregate results indicate a significant disparity, with the initiator receiving more punishment

Table 5: Punishment Points for the Initiator and Pivotal Decision-maker

		Mean punishment points				t-test
		N	Overall	Initiator	Pivotal	Initiator vs. Pivotal
Overall		162	1.42	2.30	1.82	$p = 0.030$
Default	Unequal	57	1.36	2.40	1.53	$p = 0.021$
	Equal	105	1.46	2.25	1.98	$p = 0.332$
Mechanism	Random	75	1.52	2.37	1.96	$p = 0.169$
	Grade	87	1.34	2.24	1.70	$p = 0.095$
Treatment	Random-unequal	30	1.59	2.57	1.90	$p = 0.170$
	Grade-unequal	27	1.11	2.22	1.11	$p = 0.061$
	Random-equal	45	1.48	2.24	2.00	$p = 0.528$
	Grade-equal	60	1.44	2.25	1.97	$p = 0.463$

Table notes: N = number of observations. Only unequal outcomes are considered. Initiator equals 1 if it is the first decision-maker choosing the unequal allocation. Pivotal equals 1 if it is the second decision-maker choosing the unequal allocation. P-values are from two-tailed t-tests of treatment differences, assuming equal variances.

points overall (t -test, $p = 0.030$). However, this effect is not robust across treatment variations. For example, the initiator receives significantly more punishment than the pivotal decision-maker under an unequal default, whereas the corresponding difference under an equal default is not statistically significant. In all treatment comparisons, the initiator consistently receives more punishment than the pivotal decision-maker, yet these differences generally fail to reach statistical significance.

To further examine punishment behavior, we estimate both a probit and an ordered logit model, analyzing (i) the probability of receiving punishment points and (ii) the number of points assigned conditional on punishment. This approach is motivated by the large share of zero outcomes in the data – 84.2% of observations involve no punishment (see Figure 5 in the Online Appendix A.3). The results of both models are reported in Table 6.

Results from the probit regression show that, in a specification including only mechanism and default (Model 1), the default condition emerges as a significant determinant: punishment is more likely when the default option is equal. So, unequal choices are punished more when the default is equal, i.e., when the decision maker actively chooses the unequal allocation. However, once punishment motives are included, this effect disappears. An unequal choice significantly increases the likelihood of punishment; being the initiator or the pivotal decision-maker also significantly affects the probability of being punished. The only demographic variable with a significant effect is gender: women are less likely to punish than men.

Table 6: Probit and Ordered Logit Regression

	Probit Regression				Ordered Logit Regression			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mechanism	-0.026 (0.024)				0.535 (0.314)			
Default	0.069** (0.025)				-0.313 (0.323)			
Grade-unequal		-0.029 (0.021)	-0.025 (0.030)	-0.030 (0.024)		0.786 (0.568)	0.488 (0.298)	0.571 (0.319)
Random-equal		0.065 (0.050)	0.037 (0.037)	0.051 (0.044)		-0.114 (0.319)	-0.079 (0.186)	-0.073 (0.249)
Grade-equal		0.043 (0.025)	-0.002 (0.023)	-0.015 (0.026)		0.231 (0.228)	0.475* (0.185)	0.392* (0.197)
Outcome unequal			0.028 (0.025)	0.028 (0.024)			-3.209* (1.508)	-3.203* (1.416)
Choice unequal			0.218*** (0.022)	0.218*** (0.021)			4.327*** (0.444)	4.240*** (0.410)
Outcome X Initiator			0.062** (0.020)	0.061** (0.020)			1.443 (1.304)	1.466 (1.205)
Pivotal			0.047** (0.018)	0.046** (0.018)			0.528 (1.372)	0.540 (1.271)
Female				-0.096** (0.033)				0.561* (0.283)
Age				0.004 (0.003)				0.032* (0.014)
Risk affine				0.007 (0.035)				0.038 (0.303)
Economist				-0.036 (0.039)				-0.240 (0.229)
Constant					1.390* (0.682)	1.396* (0.675)	0.166 (0.339)	0.016 (0.341)
Wald- χ^2	8.76	45.54	96.21	14777.18	3.08	3.49	127.99	512.58
$p(\chi^2)$	$p = 0.013$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p = 0.215$	$p = 0.322$	$p < 0.001$	$p < 0.001$
$p(T2 = T3)$		$p = 0.037$	$p = 0.156$	$p = 0.103$		$p = 0.139$	$p = 0.089$	$p = 0.080$
$p(T2 = T4)$		$p < 0.001$	$p = 0.482$	$p = 0.600$		$p = 0.306$	$p = 0.968$	$p = 0.595$
$p(T3 = T4)$		$p = 0.651$	$p = 0.344$	$p = 0.234$		$p = 0.196$	$p = 0.005$	$p = 0.066$
N	2430	2430	2430	2430	385	385	385	385

Table notes: Dependent variable: Punishment points from 0 to 7. Panel probit regression (Models 1 to 4) and random effect panel ordered logit regression (Models 5 to 8) with standard errors in parentheses and clustered on session. Mechanism equals 1 if grade is the group-building mechanism. Default equals 1 if the preselected option is equal. Random-unequal is the baseline category. The variables ‘Female’, ‘Economics’, ‘Risk affine’ are dummy variables for the respective socio-demographic variable. Punishment motives are dummy variables for the respective category as defined in the Online Appendix A.2 in Table 10. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

In the ordered logit model restricted to observations with at least one punishment point, neither the default condition nor the mechanism have a significant effect on punishment intensity. Unequal choice is the only significant determinant of how many punishment points are assigned, conditional on punishing. Therefore, when it comes to assigning punishment, it no longer seems to matter whether one is the initiator or the pivotal decision-maker. An interesting gender pattern emerges: while women are generally less likely to punish (Model 4), those who do punish assign significantly more punishment points than men (Model 8). Taken together, the probit and ordered logit results indicate that there are no statistically significant treatment-level differences in punishment behavior. Punishment is driven primarily by punishment motives rather than by institutional design features per se.

4.2.2 Testing for a Default Effect and the Influence of Status Differences on Punishment Behavior

We start by examining whether the default option affects responsibility attribution – specifically whether decision-makers who follow the default are punished less. Overall, 46.3% of choices coincide with the default, though patterns vary substantially across conditions. We distinguish between unequal choices implemented actively (deviating from the default) and those implemented passively (following the default).

Table 7: Testing for a Default Effect - Marginal Effects

		Initiation		Pivotality	
		No	Yes	No	Yes
Default	Unequal	0.412*** (0.069)	1.382*** (0.170)	0.465*** (0.069)	0.512** (0.170)
	Equal	0.522*** (0.071)	1.173*** (0.146)	0.532*** (0.071)	0.947*** (0.146)
Mechanism	Random	0.486*** (0.070)	1.322*** (0.158)	0.513*** (0.070)	0.945*** (0.158)
	Grade	0.445*** (0.071)	1.148*** (0.153)	0.482*** (0.071)	0.636*** (0.153)

Table notes: Average marginal effect of panel tobit regressions (see Table 13 in the Online Appendix) with standard errors in parentheses. * * * $p < 0.001$, * * $p < 0.01$, * $p < 0.01$.

To this end, Table 7 reports marginal effects from an additional regression analysis (full results are provided in Table 13 in the Online Appendix A.3), offering a more nuanced perspective on punishment behavior.

For the initiator, the results run counter to expectations. If recipients penalize active deviations from the default more harshly, punishment should be higher when the equal allocation is preselected — since choosing unequal then constitutes an active override. The data do not support this: punishment directed at initiators is slightly higher under an unequal default than under an equal

default (1.382 versus 1.173), the opposite of what the literature would predict. While the difference in absolute terms is small, statistically not significant and should be interpreted with caution, it nonetheless suggests that for initiators, the act of choosing unequal dominates any moderating effect of the default condition.

600 For the pivotal decision-maker, however, a different pattern emerges. The marginal effect associated with being pivotal is nearly halved under an unequal default compared to an equal default (0.512 versus 0.947, Wald-F-test, $p = 0.012$), indicating that unequal choices – as pivotal decision-makers – are punished considerably less severely when they coincide with the default option. This finding is consistent with the theoretical prediction that passive implementation attenuates perceived responsibility: when the pivotal decision-maker merely follows the preselected option, recipients appear to assign less blame than when the same choice requires an active override of the default. This may partly explain why, in contrast to Bartling et al. (2015), the pivotal decision-maker is punished less than the initiator in certain decision sequences.

605 Overall, punishment directed at initiators appears largely independent of the default condition, whereas the effect associated with pivotality is sensitive to it. However, once punishment motives are explicitly controlled for, the default effect largely disappears. Hypothesis 5 is therefore partially rejected.

Result 3 (Default Choice) *Choosing the default option results in less punishment when no additional punishment motives are controlled for. Selecting the default cannot shield initiators from punishment, whereas pivotal decision-makers can partially mitigate punishment by choosing the default option.*

Next, we further investigate whether punishment directed at initiators and pivotal decision-makers differs across group-formation mechanisms. The perceived legitimacy of the group-formation mechanism may shape reactions to decisions by decision-makers with different status legitimacy. Independently of the pre-study by Kraus and Traub (2026), participants evaluated the legitimacy of both mechanisms at the end of this experiment.

620 The overall legitimacy assessments remain largely unchanged between the pre-study and this experiment. If anything, the random mechanism is perceived as even more legitimate in this experiment (5.44 vs. 4.88, rating from 1 (completely not legitimate) to 7 (completely legitimate)), whereas the university entrance grade is evaluated as even less legitimate (2.57 vs. 3.09). Participants assigned through the university entrance grade perceive a random assignment as significantly more legitimate than participants assigned through the random mechanism themselves (5.82 vs. 5.09, t -test, $p < 0.001$). Detailed results are reported in Table 12 in the Online Appendix A.3.

625 Despite these pronounced differences in perceived legitimacy, we do not observe any significant effect of the allocation mechanism on assigned punishment points, either overall or specifically for initiators and pivotal decision-makers (see Table 6 and 7). While punishment for the pivotal decision-maker is slightly higher under the random mechanism in absolute terms, these differences are not statistically significant (Wald-F-test $p = 0.087$). The group-formation mechanism thus does not appear to serve as a focal point for attributing responsibility. Hypothesis 6 is therefore rejected.

630 **Result 4 (Group-building Mechanism)** *The perceived legitimacy of the group-formation mechanism does not affect punishment points assigned by recipients.*

Although legitimacy perceptions differ substantially and predictably across treatments, they do not translate into differential punishment behavior. This null finding is itself informative from a robustness perspective: it suggests that punishment heuristics such as initiation and pivotality are relatively insensitive to the legitimacy framing of role assignments, at least within the range of institutional variation studied here.

5 Discussion and Conclusion

This study set out to examine the robustness of responsibility attribution in collective decision-making by systematically extending the experimental framework introduced by Bartling et al. (2015). Our core question was a meta-scientific one: does the central finding of that study – that pivotal decision-makers receive the highest punishment – persist under theoretically meaningful variations in institutional design? By introducing default options and alternative group-formation mechanisms, we contribute to a growing literature emphasizing that replication in the social sciences should assess not only whether statistical results can be reproduced, but also whether the underlying behavioral mechanisms remain stable across contextual variations.

Our findings provide only partial support for the original result. While pivotality continues to matter for punishment behavior, we do not replicate the strong dominance of pivotal responsibility attribution reported by Bartling et al. (2015). Instead, punishment is often directed more strongly toward the initiator of an unequal allocation, particularly in situations in which the first unequal choice appears to establish the normative direction of the decision process. Moreover, default structures affect responsibility attribution for pivotal decision-makers. Surprisingly, we found no differences in punishment following a (less) legitimate group formation mechanism. Taken together, these findings suggest that responsibility attribution is not governed by a single stable heuristic, but rather by multiple overlapping mechanisms whose relative importance depends on the institutional context.

Importantly, our results should not be interpreted as a failed replication in the narrow sense. Rather, they illustrate precisely why robustness-oriented replications are crucial for cumulative scientific progress. The original findings remain informative and partially supported, yet our extensions demonstrate that seemingly (minor) procedural variations – such as the presence of defaults or changes in elicitation procedures – can meaningfully alter observed punishment patterns. This is particularly relevant for experimental economics involving social judgments, moral evaluations, and responsibility attribution, where contextual cues may shape both expectations and behavioral norms. More broadly, the study speaks directly to ongoing discussions surrounding reliability and transparency in the social sciences. Recent debates on the replication crisis have increasingly emphasized that reproducibility cannot be reduced to the binary question of whether a coefficient remains statistically significant. Reliability also concerns the extent to which behavioral relationships generalize across institutional arrangements. In this regard, our study highlights the value of systematic design variation as a complement to exact replication. By preserving the core decision environment while modifying institutional features, we are able to identify which behavioral regularities appear robust and which may be context-sensitive.

The study has several limitations that should be acknowledged transparently. First, the experiment was conducted before preregistration had become standard practice in experimental economics. While all hypotheses and analyses are reported transparently and draw on Bartling et al. (2015)

preregistration would have further strengthened the credibility of the design. Second, because
680 multiple institutional features were varied simultaneously, disentangling the exact mechanisms
behind the observed shifts in punishment behavior remains challenging. Future research could
investigate defaults, legitimacy, and elicitation procedures separately in order to isolate their inde-
pendent effects more precisely. Finally, the comparatively low frequency of unequal choices in our
study may itself have shaped punishment norms, thereby contributing to differences relative to the
685 original experiment.

Despite these limitations, the present study contributes to the meta-scientific discussion on re-
liability in two important ways. Substantively, it advances our understanding of responsibility
attribution by demonstrating that punishment behavior depends not only on pivotality, but also
on institutional framing and the sequencing of decisions. Methodologically, it illustrates how
690 replication-through-extension can generate informative evidence about the stability and boundary
conditions of experimental findings. Rather than viewing replication as a purely confirmatory ex-
ercise, our results support a more cumulative perspective in which robustness is assessed through
systematic variation and transparent comparison across designs.

References

- 695 Akbaş, M., Ariely, D., Yuksel, S., 2019. When is inequality fair? An experiment on the effect of procedural justice and agency. *Journal of Economic Behavior & Organization* 161, 114–127. doi:10.1016/j.jebo.2019.02.014.
- Albrecht, K., Essen, E.v., Fliessbach, K., Falk, A., 2013. The influence of status on satisfaction with relative rewards. *Frontiers in Psychology* 4. doi:10.3389/fpsyg.2013.00804.
- 700 Ball, S., Eckel, C.C., 1998. The economic value of status. *Journal of Socio-Economics* 27, 495–514.
- Bartling, B., Fischbacher, U., Schudy, S., 2015. Pivotality and responsibility attribution in sequential voting. *Journal of Public Economics* 128, 133–139. doi:10.1016/j.jpubeco.2015.03.010.
- 705 Berry, J., Coffman, L.C., Hanley, D., Gihleb, R., Wilson, A.J., 2017. Assessing the rate of replication in economics. *American Economic Review* 107, 27–31. doi:10.1257/aer.p20171119.
- Bhatia, D., Fischbacher, U., Hausfeld, J., Stumpf, R., 2024. Blame and praise: Responsibility attribution patterns in decision chains. *Experimental Economics* 27, 637–663.
- 710 Blue, P.R., Hu, J., Wang, X., van Dijk, E., Zhou, X., 2016. When do low status individuals accept less? the interaction between self- and other-status during resource distribution. *Frontiers in Psychology* 7. doi:10.3389/fpsyg.2016.01667.
- Bock, O., Baetge, I., Nicklisch, A., 2014. hroot: Hamburg registration and organization online tool. *European Economic Review* 71, 117–120. doi:10.1016/j.euroecorev.2014.07.003.
- 715 Brandts, J., Charness, G., 2000. Hot vs. cold: Sequential responses and preference stability in experimental games. *Experimental Economics* 2, 227–238. doi:10.1023/a:1009962612354.
- Camerer, C.F., Dreber, A., Forsell, E., Ho, T.H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., Wu, H., 2016. Evaluating replicability of laboratory experiments in economics. *Science* 351, 1433–1436. doi:10.1126/science.aaf0918.
- 720 Caputo, R., Duch, R., 2024. Pivotal decision maker, agenda power and collective responsibility attribution .
- 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:10.1016/j.jbef.2015.12.001.
- 725 Chockler, H., Halpern, J.Y., 2004. Responsibility and blame: A structural-model approach. *Journal of Artificial Intelligence Research* 22, 93–115.

- 730 Collaboration, O.S., 2015. Estimating the reproducibility of psychological science. *Science* 349, aac4716.
- DeScioli, P., Bruening, R., Kurzban, R., 2011a. The omission effect in moral cognition: Toward a functional explanation. *Evolution and Human Behavior* 32, 204–215. doi:10.1016/j.evolhumbehav.2011.01.003.
- 735 DeScioli, P., Christner, J., Kurzban, R., 2011b. The omission strategy. *Psychological Science* 22, 442–446. doi:10.1177/0956797611400616.
- Duch, R., Przepiorka, W., Stevenson, R., 2014. Responsibility attribution for collective decision makers. *American Journal of Political Science* 59, 372–389. doi:10.1111/ajps.12140.
- Engl, F., 2018. A theory of causal responsibility attribution. SSRN working paper 2932769 .
- 740 Fehr, E., Gintis, H., 2007. Human motivation and social cooperation: Experimental and analytical foundations. *Annual Review of Sociology* 33, 43–64. doi:10.1146/annurev.soc.33.040406.131812.
- Feltovich, N., 2019. Is earned bargaining power more fully exploited? *Journal of Economic Behavior & Organization* 167, 152–180. doi:10.1016/j.jebo.2019.09.021.
- 745 Fershtman, C., Gneezy, U., List, J.A., 2012. Equity aversion: Social norms and the desire to be ahead. *American Economic Journal: Microeconomics* 4, 131–144. doi:10.1257/mic.4.4.131.
- Fischbacher, U., Schudy, S., 2020. Agenda control and reciprocity in sequential voting decisions. *Economic Inquiry* 58, 1813–1829. doi:10.1111/ecin.12898.
- 750 Fleiß, J., 2015. Merit norms in the ultimatum game: An experimental study of the effect of merit on individual behavior and aggregate outcomes. *Central European Journal of Operations Research* 23, 389–406. doi:10.1007/s10100-015-0385-8.
- Gärtner, M., 2018. The prosociality of intuitive decisions depends on the status quo. *Journal of Behavioral and Experimental Economics* 74, 127–138. doi:10.1016/j.socec.2018.04.005.
- 755 Gärtner, M., Sandberg, A., Yechiam, E., 2017. Is there an omission effect in prosocial behavior? A laboratory experiment on passive vs. active generosity. *PloS one* 12, e0172496. doi:10.1371/journal.pone.0172496.
- Grimm, S., Klimm, F., 2019. Blaming the refugees? experimental evidence on responsibility attribution. *Journal of Economic Psychology* 72, 156–178. doi:https://doi.org/10.1016/j.joep.2019.03.004.
- 760 Henrich, J., Gil-White, F.J., 2001. The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior* 22, 165–196. doi:10.1016/s1090-5138(00)00071-4.

- Jachimowicz, J.M., Duncan, S., Weber, E.U., Johnson, E.J., 2019. When and why defaults influence decisions: A meta-analysis of default effects. *Behavioural Public Policy* 3, 159–186. doi:10.1017/bpp.2018.43.
- 765
- Johnson, E.J., Goldstein, D., 2003. Do defaults save lives? *Science* 302, 1338–1339. doi:10.1126/science.1091721.
- Johnson, E.J., Hershey, J., Meszaros, J., Kunreuther, H., 1993. Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty* 7, 35–51. doi:10.1007/bf01065313.
- 770
- Kraus, J., Traub, S., 2026. On the descriptive legitimacy of status differences. Working paper doi:10.48548/pubdata-3703.
- Levitt, S.D., List, J.A., 2009. Field experiments in economics: The past, the present, and the future. *European Economic Review* 53, 1–18. doi:10.1016/j.euroecorev.2008.12.001.
- 775
- Maniadis, Z., Tufano, F., List, J.A., 2017. To replicate or not to replicate? Exploring reproducibility in economics through the lens of a model and a pilot study. *The Economic Journal* 127, 209–235. doi:10.1111/eoj.12527.
- Ritov, I., Baron, J., 1992. Status-quo and omission bias. *Journal of Risk and Uncertainty* 5, 49–61.
- 780
- Samuelson, W., Zeckhauser, R., 1988. Status quo bias in decision making. *Journal of Risk and Uncertainty* 1, 7–59. doi:10.1007/bf00055564.
- Shultz, T.R., Schleifer, M., Altman, I., 1981. Judgments of causation, responsibility, and punishment in cases of harm-doing. *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement* 13, 238–253. doi:10.1037/h0081183.
- 785
- Spellman, B.A., 1997. Crediting causality. *Journal of Experimental Psychology: General* 126, 323–348. doi:10.1037/0096-3445.126.4.323.
- Spranca, M., Minsk, E., Baron, J., 1991. Omission and commission in judgment and choice. *Journal of Experimental Social Psychology* 27, 76–105. doi:10.1016/0022-1031(91)90011-t.
- 790
- Sunstein, C.R., Thaler, R.H., 2003. Libertarian paternalism is not an oxymoron. *The University of Chicago Law Review* 70, 1159. doi:10.2307/1600573.
- Tyler, T.R., 2006. Psychological perspectives on legitimacy and legitimation. *Annual Review of Psychology* 57, 375–400. doi:10.1146/annurev.psych.57.102904.190038.

Online Appendix: Responsibility Attribution Under Institutional Variation: A Robustness-Oriented Replication in Sequential Group Decisions

795

Janina Kraus

A Further Results

A.1 Sample

Table 8: Breakdown of the Sample

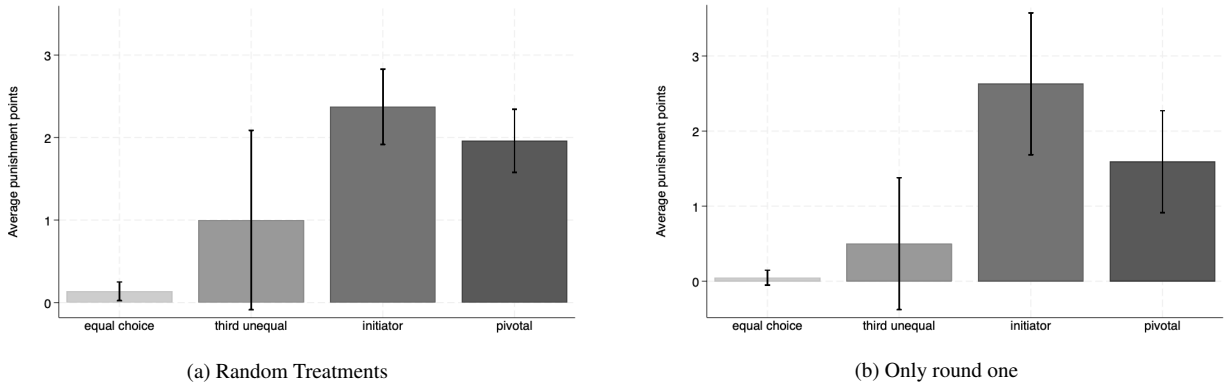
	N	Female	Age	Semester	Mean grade
Random-unequal (T1)	72	66.67%	26.23	6.71	2.02
Grade-unequal (T2)	66	62.12%	26.86	9.12	2.00
Random-equal (T3)	66	62.50%	26.99	8.69	2.08
Grade-equal (T4)	66	69.70%	27.00	9.03	2.12
Overall	270	65.19%	26.77	8.40	2.06

800 **A.2 Comparison to Bartling et al. (2015)**

Table 9: Comparison to Bartling et al. (2015) - Robustnesscheck

Allocation	Decision sequence	Decision-maker 1	Decision-maker 2	Decision-maker 3	Pivotal vs. all	Pivotal vs. initiator	Initiator vs. all
		Only random			Wilcoxon rank-sum		
Unequal	u-u-u	1.11	0.89	1.00	$p = 0.977$	$p = 0.920$	$p = 0.885$
	u-u-e	2.42	2.25	0.17	$p = 0.033$	$p = 0.617$	$p = 0.006$
	u-e-u	2.59	0.19	2.26	$p = 0.021$	$p = 0.654$	$p = 0.003$
	e-u-u	0.00	2.67	1.60	$p = 0.271$	$p = 0.154$	$p = 0.004$
Equal	u-e-e	1.36	0.14	0.19			
	e-u-e	0.10	1.31	0.10			
	e-e-u	0.05	0.05	2.57			
	e-e-e	0.17	0.12	0.16			
		Only round one			Wilcoxon rank-sum		
Unequal	u-u-u	1.33	0.50	0.50	$p = 0.823$	$p = 0.703$	$p = 0.655$
	u-u-e	2.22	1.44	0.11	$p = 0.461$	$p = 0.521$	$p = 0.070$
	u-e-u	3.22	0.00	2.22	$p = 0.304$	$p = 0.443$	$p = 0.026$
	e-u-u	0.00	4.67	2.33	$p = 0.892$	$p = 0.261$	$p = 0.057$
Equal	u-e-e	2.50	0.00	0.00			
	e-u-e	0.00	1.93	0.00			
	e-e-u	0.00	0.00	3.00			
	e-e-e	0.21	0.15	0.27			

Table notes: “u” denotes a choice of the unequal allocation; “e” denotes a choice of the equal allocation. The three rightmost columns show p-values of of rank-sum test comparing the punishment for the pivotal decision-maker to the punishment for the two other decision-makers (“pivotal vs. al”) and to the punishment for the initiator only (“pivotal vs. initiator”). Additionally, a comparison between the punishment for the initiator to the punishment for the two other decision-makers (“initiator vs. all“) is added. Only round one: p-values should be interpreted with caution, as some cells contain only a few observations.



Notes: Mean of assigned punishment points (from 0 to 7) for the decision-makers. The two random treatments are considered on the left, while the punishment decisions from the first round are considered on the right. *Equal choice* means choosing the equal allocation if the two others choose unequal. *Third unequal* represents the third decision-maker choosing unequal if the two others also chose unequal. *Initiation* means the first decision-maker choosing the unequal allocation. *Pivotality* means the pivotal decision-maker finally determining the unequal allocation. Bars represent the 95% confidence interval.

Figure 3: Average Punishment If the Unequal Allocation Results - Robustnesscheck

Table 10 illustrates how the decision-makers correspond to the different punishment motives used in this study.

Table 10: Categorization of Decision-makers

Sequence	Decision-maker 1	Decision-maker 2	Decision-maker 3	
u-u-u	(2), (3), (6)	(2), (3), (7)	(2), (5)	} (1)
u-u-e	(2), (3), (6)	(2), (3), (7)	(4)	
u-e-u	(2), (3), (6)	(4)	(2), (3), (7)	
e-u-u	(4)	(2), (3), (6)	(2), (3), (7)	
u-e-e	(2), (3), (6)	(4)	(4)	
e-u-e	(4)	(2), (3), (6)	(4)	
e-e-u	(4)	(4)	(2)	
e-e-e	(4)	(4)	(4)	

Notes: The table shows which decision-makers are placed into the seven categories defined above for the different sequences. For example, the top row shows the sequence in which all decision-maker chose the unequal allocation (u-u-u). Thus, all three decision-maker are classified into the category (1) unequal outcome and (2) unequal choice. Additionally, decision-maker 1 and 2 are classified as unkind (category (3)), as their choices are still impactful. The first decision-maker is also the one starting the sequence and thus, categorized as (6) the initiator while the second decision-maker is pivotal for the unequal outcome (7). Furthermore, the third decision-maker falls into category (5) as he chooses unequal although the unequal allocation is already implemented.

(1) Outcome unequal:

805 If the unequal allocation is implemented, the decision-makers are sorted into this category, regardless of their choice. This motive refers to the group-level, while the individual choices determine the following motives.

(2) Unequal choice:

Whenever a decision-maker chooses unequal, he is placed in this category.

(3) Unkindness:

810 In contrast to ‘unequal choice’, a decision-maker is classified as unkind as long as his decision has an impact on the implementation of the unequal outcome. Relatedly, an unequal choice increases the probability of implementing the unequal outcome. Consequently, a decision-maker who chooses unequally is classified as having an unkind intention if no majority is reached.

(4) Equal choice:

815 All decision-makers who choose the equal allocation are placed in this category.

(5) Third unequal:

A decision-maker is placed in this category if the unequal outcome is already implemented (the first and second decision-makers choose unequally) and the third decision-maker also

820 chooses the unequal allocation. This category exists only once, namely in the first decision
sequence (u-u-u).

(6) Initiation:

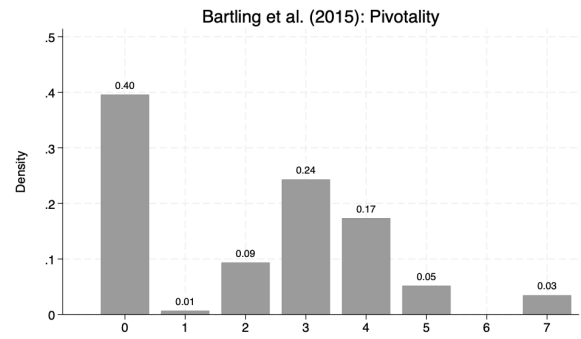
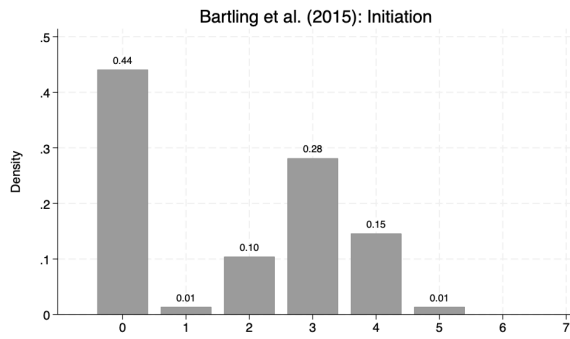
825 The first decision-maker choosing the unequal allocation is classified as initiator. Addition-
ally, any decision-maker who chooses unequally can be considered to have an unkind inten-
tion, as long as he is influential.

(7) Pivotality:

830 The decision-maker who determines the outcome is the pivotal decision-maker. Thus, an-
other decision-maker (either the first or the second) has already chosen the unequal alloca-
tion, so the pivotal decision-maker can still influence and determine the outcome in both
directions, either to an equal or unequal outcome.

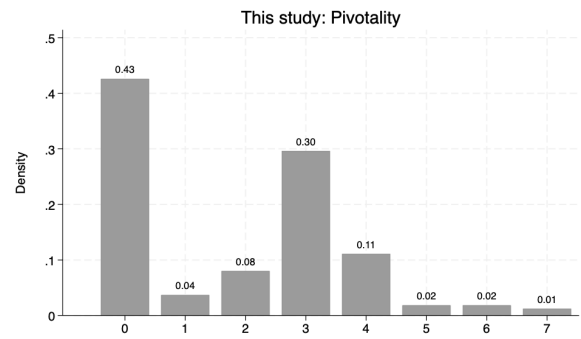
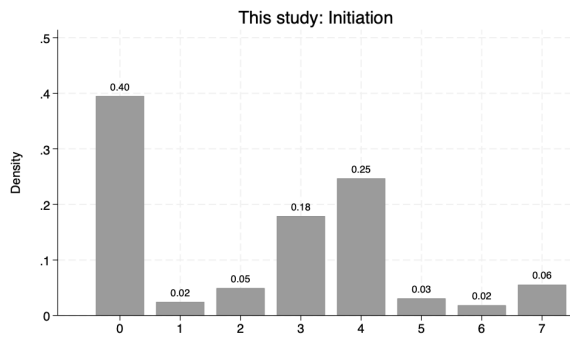
Of course, there are different ways to define punishment motives. Because they often overlap, some
decision-makers are placed in more than one category. Whenever a decision-maker is classified as
pivotal, it is obvious that he also fits into the category of unkindness and unequal choice, since
he necessarily chose unequal to be pivotal. Furthermore, unkindness can be defined in two ways:
835 either every unequal allocation decision is classified as unkind, or only decisions that affect the
outcome are classified as unkind, which is the case in this study.

Additionally, in Bartling's experiment, the sixth category, "initiation", did not exist in that sense.
However, since it makes sense to make a further distinction between "unkind intention" (which in-
cludes pivotal decision-makers and initiators) and "pivotality". What is unique about the distinction
840 between initiation and pivotality is that these two roles are mutually exclusive. A decision-maker
is either the initiator or the pivotal decision-maker, but cannot be both. This strict distinction was
not present in the original study by Bartling et al. (2015).



(a)

(b)



(c)

(d)

Notes: The figure shows the distribution of assigned punishment points (ranging from 0 to 7) between the initiator and pivotal decision-maker.

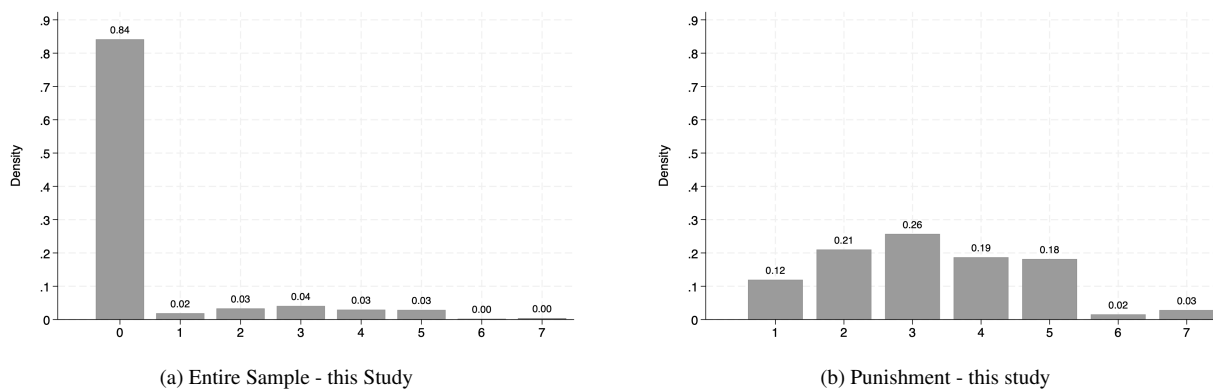
Figure 4: Distribution of Punishment Points for the Initiator and Pivotal Decision-maker

Table 11: Comparison of Different Punishment Motives

	Only round one	Only random treatments
Outcome unequal	-0.119 (0.073)	-0.004 (0.088)
Unequal choice	0.452 (0.318)	0.864 (0.482)
Initiator	1.648** (0.553)	0.577 (0.574)
Outcome unequal X Initiator	0.482 (0.647)	0.796 (0.422)
Pivotal	1.093** (0.403)	0.960 (0.482)
Constant	0.167** (0.056)	0.140* (0.059)
R^2	0.350	0.330
N	405	1242

Table notes: Random effects OLS regression with punishment points for decision-makers as dependent variable (from 0 to 7). All punishment variables are dummy variables. *Outcome unequal* equals 1 if the unequal allocation is implemented. *Unequal choice* equals 1 if the unequal allocation is chosen. *Intention unkind* equals 1 if a decision-maker opts for the unequal allocation without having a majority. *Initiation* equals 1 if a decision-maker is the first one in a decision sequence choosing the unequal outcome. *Pivotal* equals 1 if a decision-maker is pivotal for the unequal outcome. Standard errors are in parentheses and clustered on 270 individuals in this study and 72 individuals in Bartling et al. (2015). *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

A.3 Treatment Comparison



Notes: Punishment points range from 0 to 7. On the left, the entire sample is considered. Only observations where at least one punishment point is assigned are considered on the right side.

Figure 5: Histogram of Punishment Points

Table 12: Rating and Deservingness of Group-Building Mechanism

		Rating random	Rating grade	Deserving role A	Deserving role B
Pretest		4.88	3.09	/	/
Experiment		5.44	2.57	5.72	5.02
Mechanism	Random (T1+T2)	5.09	2.46	5.83	5.23
	Grade (T3+T4)	5.82	2.68	5.61	4.80
<i>Assigned Role</i>		<i>Mechanism</i>			
Decision-Maker	Random	5.01	2.55	5.71	4.93
	Grade	5.71	2.85	6.23	4.48
Recipient	Random	5.16	2.36	5.94	5.54
	Grade	5.92	2.52	5.00	5.12

Table notes: The rating scale ranges from 1 (completely not legitimate) to 7 (completely legitimate). The deservingness scale ranges from 1 (completely not deserving) to 10 (completely deserving).

Table 13: Testing for a Default and Mechanism Effect

	(7)	(8)	(9)	(10)
Mechanism	-0.048 (0.100)	-0.051 (0.100)		
Default			0.090 (0.098)	0.091 (0.098)
Initiator		0.767*** (0.125)		0.764*** (0.125)
Pivotal	0.282* (0.125)		0.283* (0.125)	
Unequal#Initiator	0.860*** (0.187)			
Unequal#Pivotal		-0.020 (0.187)		
Random#Initiator			0.877*** (0.177)	
Random#Pivotal				0.463** (0.177)
Outcome unequal	-0.083 (0.082)	-0.072 (0.082)	-0.079 (0.082)	-0.080 (0.082)
Choice unequal	1.533*** (0.073)	1.529*** (0.073)	1.532*** (0.073)	1.533*** (0.073)
Female	-0.265* (0.105)	-0.261* (0.105)	-0.264* (0.105)	-0.264* (0.105)
Age	0.016 (0.009)	0.016 (0.009)	0.016 (0.009)	0.016 (0.009)
Risk affine	-0.003 (0.118)	-0.000 (0.118)	-0.002 (0.118)	-0.002 (0.118)
Economist	-0.201 (0.124)	-0.200 (0.124)	-0.200 (0.124)	-0.199 (0.124)
Constant	-0.017 (0.262)	-0.037 (0.262)	-0.162 (0.277)	-0.157 (0.277)
N	2430	2430	2430	2430

Table notes: Dependent variable: Punishment points (from 0 to 7). Random effects tobit regression with standard errors in parentheses. Mechanism equals 1 if grade is the group-building mechanism. Default equals 1 if the preselected option is equal. Random-unequal is the baseline category. The variables ‘Female’, ‘Economics’, ‘Risk affine’ are dummy variables for the respective socio-demographic variable. Punishment motives are dummy variables for the respective category as defined in the Online Appendix A.2 in Table 10. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

B Instructions

845 **General explanations for the participants**

Welcome to the experiment. If you read the instructions carefully and follow the rules, you can earn money in this experiment. You will receive **6 euros as a fixed payout**. Depending on your decisions and the decisions of the other participants, you can earn **additional money**. The money will be paid to you in cash immediately afterwards. During the experiment we do not speak of
850 euros, but of points. These are converted according to the following exchange rate:

1 point = 1 euro

Talking to other participants is not allowed during the entire experiment. If you have any questions, please direct them only to us. We will be happy to answer your questions individually. It is very important that you follow this rule. Otherwise, the results of this experiment will be scientifically
855 worthless.

This experiment consists of **2 parts**, both parts will be explained one after the other. The experiment should take 60 to 75 minutes to complete. The first part of the experiment is explained below.

Detailed information about the first part of the experiment

860 The first part of the experiment is about the distribution of 10 points between two participants A and B. While participant A decides about the distribution of the points, participant B has the possibility to take points away from participant A.

You will first take the role of participant A and then the role of participant B. This means that you will first have the opportunity to determine the distribution of points and then to distribute the
865 deduction points.

Your task as “Participant A”:

First, you are participant A and you choose how the 10 points will be divided between you and a random other participant B. You have to choose between two possible allocations: allocation 1 (5;
870 5), where you (participant A) and participant B each get 5 points, and allocation 2 (9; 1), where you (participant A) get 9 points and participant B gets 1 point. In Figure 6 you see your decision screen as participant A.

Participant A - Allocation

First, you are participant A and have to choose between **allocation 1 (5; 5)** and **allocation 2 (9; 1)**.

Which allocation do you choose?

- Allocation 1 (5; 5)
 Allocation 2 (9; 1)

Figure 6: Decision Screen as Participant A

Your task as “Participant B”:

875 Afterwards you decide as participant B. Since you are now in the position of participant B, you can assign deduction points to participant A. In Figure 7 you can see the two possible allocations 1 and 2, where you can assign up to 5 deduction points to participant A for each allocation.

Participant B - Deduction Points

You are now in the position of participant B and can assign deduction points to participant A. You can assign **up to 5 deduction points** to participant A for **each allocation**.

Assume participant A has chosen allocation 1 (5; 5) .	How many deduction points do you want to distribute to participant A?
	<input type="text"/> points
Assume participant A has chosen allocation 2 (9; 1) .	How many deduction points do you want to distribute to participant A?
	<input type="text"/> points

If you do not want to deduct any points from an allocation, enter 0 in the respective field.

Figure 7: Decision Screen as Participant B

Deducting points has a **cost**. If you want to deduct points from participant A, you must give up 1 point to be able to deduct up to 5 points. **Any integer number of points between 0 and 5 points** can be deducted for both possible choices. Once at least 1 point is subtracted, the cost is 1 point. 880 So the cost of subtracting points is always 1 point, regardless of the number of points subtracted.

Your payout from part 1:

After you have made your decision as participant A and participant B, you will be randomly assigned to another participant in the experiment. Both of you will then be randomly assigned a role, 885 either participant A or participant B, and your decision in that role will be selected as **relevant for payout**.

Your payout as “Participant A”

If you are selected as participant A, you will receive the number of points you kept for yourself 890 minus the number of points deducted by your assigned participant B. For example, if you selected allocation 1 (5; 5) and participant B deducted 2 points from your allocation, you would receive 5 points - 2 points = 3 points.

Your payout as “Participant B”

If you are selected as participant B, you will receive the number of points that participant A has 895 assigned to you, minus 1 point if you have assigned deductions for that assignment. That is, if contestant A chose allocation 1 (5; 5) and you gave away 2 points for that allocation, you will receive 5 points - 1 point = 4 points. If you did not deduct any points for allocation 1 (5; 5), you do not incur a cost of 1 point, regardless of whether and how many points you deducted for allocation 900 2 (9; 1).

You will know your payout from this part of the experiment at the end of the entire experiment. Before we proceed with the second part of the experiment, we will ask you to give us your high school graduation grade. Then we will give you the instructions for the second part of the experi-

905 ment.

If you have any questions, raise your hand outside the cabin and wait quietly until someone comes to you. Thank you for your participation and have fun with the experiment.

910 **Instructions for the second part of the experiment**

Your payout from the first part of the experiment is now fixed and will be sent to you at the end of the experiment. The following second part of the experiment has no effect on your payout from the first part. You will now receive the **instructions for the second part of the experiment**.

915

Role assignment in part 2:

At the beginning of the second part, you will be assigned to a role of participant A or participant B *randomly* [based on your final grade. Those with the relatively higher grades will be assigned to participant A and the others to participant B.] This role assignment will be maintained throughout
920 the second part of the experiment.

Part 2 of the experiment consists of a total of 6 rounds. At the beginning of each new round, you will be randomly and anonymously divided into groups with other participants A and B who are also taking part in this experiment. Each group consists of three participants A (A1, A2, and A3)
925 and three participants B (B1, B2, and B3). Thus, you will always be paired with two other participants in your role and three other participants in the other role. All participants remain anonymous, as do you. After each round, the groups are rearranged, with each participant keeping his or her role A or B. However, you will be grouped with different participants A and B than in the previous round.

930

The decision-making situation:

The second part of the experiment consists of two stages. In **stage 1**, the three participants A of a group decide **sequentially** how to divide 30 points between the three participants A and the three participants B within their group. Participants A must choose between two possible allocations of
935 points:

- **Allocation 1 (5, 5, 5; 5, 5, 5):** Participant A and participant B receive 5 points each.
- **Allocation 2 (9, 9, 9; 1, 1, 1):** The three participants A receive 9 points each and the three participants B receive 1 point each.

The allocation that receives the majority of votes from participant A will be implemented. Then, in
940 **stage 2**, participants B will have the opportunity to assign deduction points to the three participants A.

Stage 1 - Allocation decision:

At the beginning of each new round, participants A learn the position (A1, A2 or A3) to which
945 they have been assigned for that round. It is possible to get the same position several times, even consecutively. However, the group composition changes with each new round. The participants A

vote **sequentially** on the allocation, whereby the allocation 1 (5, 5, 5; 5, 5, 5) [2 (9, 9, 9; 1, 1, 1)] is already preselected. Each participant A then has 30 seconds to choose one of the two allocations.

950 **The decision as “Participant A1”:**

The first participant to make a decision is participant A1, who has 30 seconds to select an allocation and click the ‘next button’. If contestant A1 has not clicked the ‘next button’ before the time expires, the marked allocation will be scored.

Round 1 from 6 - Participant A1

Time to make a decision: **0:21**

Which allocation do you choose?

Allocation 1 (5, 5, 5; 5, 5, 5) Allocation 2 (9, 9, 9; 1, 1, 1)

Next

Figure 8: Decision Screen of Participant A1

The decision as “Participant A2”:

955 Participant A, who makes the second decision, is participant A2. Participant A2 is shown how participant A1 decided before making his own decision. Again, an allocation is preselected and participant A2’s decision screen is displayed for 30 seconds.

The decision as “Participant A3”:

960 Participant A, who makes the third decision, becomes participant A3. Participant A3 is shown how participants A1 and A2 have decided before being given 30 seconds to make his own decision.

The distribution for which **at least two of the three participants A decide** is implemented for that round. The voting result is therefore fixed as soon as two participants A have chosen the same distribution.

965

The decision as “Participant B”:

While participants A make their decisions one by one, we would like to know from participants B which division they think participants A1, A2 and A3 will choose. For this purpose, participants B are assigned to one of the three positions B1, B2 and B3 in each new round. They then go through the decision situation one by one, just like participants A, and have 30 seconds to state their expected distribution. This means that participant B1 states his expectations first, followed by participant B2, who is shown the expectation stated by participant B1. After that, participant

970

975 B3 states his expected distribution, while being shown the expectations of participants B1 and B2. (see Figure 9).

Round 1 from 6 - Participant B3

Time to make a decision: **0:17**

Assume participant A1 has chosen **allocation 1 (5, 5, 5; 5, 5, 5)** .
Assume participant A2 has chosen **allocation 1 (5, 5, 5; 5, 5, 5)** .

Which allocation do you expect from participant A3?

Allocation 1 (5, 5, 5; 5, 5, 5) Allocation 2 (9, 9, 9; 1, 1, 1)

Next

Figure 9: Decision Screen of Participant B3

Stage 2 – Distribution of deduction points:

After participants A have determined their allocation and participants B have stated their expectations, the next step is for participants B to distribute their deduction points.

980 **The decisions of participants B:**

Participant B learns the allocation that has been made and how each participant A has decided. **Each participant B** then has the opportunity to deduct points from the payouts of participants A1, A2 and A3.

985 Deducting points has a **cost**, as in part 1 of the experiment: If participant B wants to deduct points from participant A, he must give up 1 point in order to be able to deduct up to 7 points from participant A in his group. **Any integer number of points between 0 and 7 points can be deducted.** Once at least 1 point has been deducted, the participant B who deducted the points has a cost of 1 point. The cost of deducting points is therefore always 1 point, regardless of the number of points deducted. For example, if participant B wants to deduct 7 points from participant A3, participant
990 A3's payout is reduced by **7 points** and participant B's payout is reduced by **1 point**.

The only restriction on the deduction of points is that no more than a total of 7 points can be deducted and that no more points can be deducted from a participant A than he has received in the implemented allocation. If the allocation is 1 (5, 5, 5; 5, 5, 5), a total of up to 7 points can
995 be deducted, but no participant A can be deducted more than 5 points.

At the end of the six rounds, one round is randomly selected as relevant to the payout. In each group, **a participant B is then randomly selected** and only the deduction points of this one

randomly selected participant B are converted. In each group, therefore, **only this one participant B incurs a cost of 1 point in this one round** (if he deducts points).

Since you do not know which round will be selected for payout and whether you will be drawn, the deduction points you distribute in each of the 6 rounds may determine the payouts at the end of the experiment.

The decision as “Participant A”:

While participants B are assigning their deduction points, we want to know from participants A how many deduction points they expect to receive. Thus, as participant A, you will see which allocation was chosen by participants A1, A2, A3 within your group, and you must indicate how many deduction points you think participant A will receive for the decision he made.

Your payoff from part 2 of the experiment:

After the 6th round you will see which round and which anonymous participant B was chosen as relevant for the payout. You will also see which distribution was chosen by the majority in the payout relevant round and if and how many deduction points were distributed by the randomly chosen participant B.

Your payout as “Participant A”:

As participant A, your payout will be equal to your allocation minus the deduction points of the randomly selected participant B. For example, if the majority of your group chose allocation 1 (5, 5, 5; 5, 5, 5), and the selected participant B distributed 1 deduction point to each participant A, you as participant A will receive **5 points minus 1 deduction point, so 4 points** from part 2.

Your payout as “Participant B”:

As a participant B, your payout will be equal to your assigned allocation minus one point if you are selected to distribute deductions.

If the majority of participants A in your group chose allocation 1 (5, 5, 5; 5, 5, 5), you will receive 5 points. If you are also participant B, who is allowed to assign deduction points, and you have done so, you will be deducted 1 point. So you get **5 points minus 1 point (cost of the deduction points), i.e. 4 points** from part 2.

After the second part, the actual experiment is over. We will ask you to answer the following questionnaire. You will then receive your payout. Please remain seated until we call your cubicle number.

If you have any questions, please raise your hand outside the booth and wait quietly until someone comes to you. If you have no questions, you will be given the password to start the second part of the experiment.

**STUDY: RESPONSIBILITY ATTRIBUTED UNDER INSTITUTIONAL VARIATION:
A ROBUSTNESS-ORIENTED REPLICATION IN SEQUENTIAL GROUP DECISIONS**

DOI <https://doi.org/10.48548/pubdata-3832>

Research Cluster Working Paper Series “Behavioral Economics and Societal Transformation”
ISSN 3055-1897

Version dated: 15.06.2026

This working paper was produced within the context of the [Research Cluster “Behavioral Economics and Social Transformation”](#), a collaboration between seven universities in Lower Saxony (Clausthal University of Technology, Georg-August University of Göttingen, Leibniz University Hannover, Leuphana University Lüneburg, Carl von Ossietzky University of Oldenburg, University of Osnabrück, University of Vechta) and the Fraunhofer Institute for Applied Information Technology. As part of the “Wissenschaftsräume” funding line, the project is supported by the Volkswagen Foundation and the Lower Saxony Ministry of Science and Culture through the [zukunft.niedersachsen](#) program. Further information on the project and additional working papers can be found at www.wissenschaftsraum-verhalten.de.

Contact:
Research Cluster “Behavioral Economics and Social Transformation”
Fon +49 4134 667 2938
Mail wissenschaftsraum-verhalten@leuphana.de

Dr. Janina Kraus
janina.kraus@leuphana.de

© 2026 This working paper is licensed under the Creative-Commons-License CC BY 4.0 license. For the full text of the licence terms, please visit the URL <https://creativecommons.org/licenses/by/4.0/>

Institute of Economics (IVWL)
Leuphana University Lüneburg
Universitätsallee 1
21335 Lüneburg