No. 126

THURGAU INSTITUTE OF ECONOMICS
at the University of Konstanz

R. Anselm, D. Bhatia,
U. Fischbacher,
J. Hausfeld

Blame and Praise: Responsibility Attribution Patterns in Decision Chains

Research Paper Series
Thurgau Institute of Economics
and Department of Economics
at the University of Konstanz

Blame and Praise: Responsibility Attribution Patterns in Decision Chains *

Regina Anselm †1, Deepti Bhatia[1], Urs Fischbacher[1], and Jan Hausfeld[2]

[1]Department of Economics, University of Konstanz, and Thurgau Institute of Economics
[2]CREED and Amsterdam School of Economics, University of Amsterdam

March 15, 2022

Abstract

How do people attribute responsibility when an outcome is not caused by a single person but results from a decision chain involving several people? We study this question in an experiment, in which five voters sequentially decide on how to distribute money between them and five recipients. The recipients can reward or punish each voter, which measures responsibility attribution. In the aggregate, we find that responsibility is attributed mostly according to the voters' choices and the pivotality of the decision, but not for being the initial voter. On the individual level, we find substantial heterogeneity with three overall patterns: Little to no responsibility attribution, pivotality-driven, and focus on choices. These patterns are similar when praising voters for good outcomes and blaming voters for bad outcomes.

Keywords: Responsibility Attribution, Collective Decision-Making, Voting, Decision Process

JEL Classification: C91, C92, D63, D70, D91

*We thank the audience at the GfeW 2021 and the research group of the Thurgau Institute of Economics for fruitful discussions and helpful comments. We gratefully acknowledge support from the German Research Foundation (DFG) through research unit FOR 1882 "Psychoeconomics".

†Corresponding author.
E-mail addresses:
regina.anselm@uni-konstanz.de, deepti.bhatia@uni-konstanz.de
urs.fischbacher@uni-konstanz.de, j.hausfeld@uva.nl

1 Introduction

How is responsibility attributed when an outcome results from a chain of actions? Take penalty shootouts in soccer for example: Who is held responsible for the defeat or the victory? Does the order matter, i.e., is the kicker who misses first or the kicker who misses last blamed more for the defeat? Similar questions also arise when disasters happen (i.e., sinking of big ships) since on a larger scale, disasters are often the results of a chain of unfortunate circumstances, decisions, and actions. Whittingham (2004) presents several examples of disasters and discusses the responsibility of the different people involved. Typically, someone makes a mistake which is not detected or appropriately fixed. This mistake causes/adds to further problems until a disaster is unavoidable. Similarly, good outcomes are also often the result of the (sequential) interaction of people – in a penalty shootout there is also always a winning team. One important example of a positive decision chain is joint production. If a firm releases a product, research and development, production, and marketing sequentially contribute to the success. In our study, we experimentally investigate how people attribute responsibility in decision chains by allowing people to allocate blame and praise to others.

The general question of responsibility attribution has been addressed from different angles. A normative point of view has been taken from a philosophical (Feinberg, 1970) as well as from a legal perspective (Hart and Gardner, 2008). More recently, the question has also attracted the interest of psychologists (Ross and Nisbett, 1991; Weiner, 1995; Gerstenberg et al., 2011) political scientists (Iyengar, 1994) and economists, both from an empirical (Charness, 2000; Bartling and Fischbacher, 2012; Bartling et al., 2015; Duch et al., 2015) and from a theoretical perspective (Besley, 2006; Bartling and Fischbacher, 2012; Engl, 2018). Our study has an empirical point of view as we investigate in a lab experiment how people assign responsibility. Understanding the empirical patterns of responsibility attribution is important because it has consequences for how we setup liability rules and how we distribute the benefits of joint ventures. Our experiment shows what rules of responsibility attribution people spontaneously apply, and the analysis of the heterogeneity provides insights on how well people agree with respect to these rules.

In real life decision chains, the decision makers and the actions differ in many dimensions. In our experiment, we study the impact of the sequence in isolation. For this reason, we investigate decisions in a sequential voting game, in which symmetric voters decide with majority over a good or bad outcome for other people. The subjects in our experiment are matched into groups of five voters and five recipients. The five voters choose sequentially between two options of how to allocate points between voters and recipients, while the outcome is determined by simple majority rule. The two allocations differ in their fairness: the unfair allocation favors the voters, while the fair allocation results in similar payoffs for the voters and recipients. The recipients receive full information about the voting sequence and then have the possibility to sanction the voters. Our treatments differ in the sanctioning options. There is a treatment with only punishment, one with only reward, and one with both. Finally, we use process measures and record the response times for all participants and use eye-tracking for the recipients.

We make several contributions to the existing literature. We study different outcomes (good and bad outcomes), compare different sanctioning options (punishment and reward), explore more motives (outcome, choice, intention, initiation, pivotality, causal responsibility based on models), and add process measures (response time and eye-tracking). We explore the following research questions:

First, we study *how responsibility is attributed to different roles in the decision chain.* Two agents are in a particular focus, the initiator (first voter in favor of the resulting outcome) and the pivotal decision maker (in our case the third voter in favor of the resulting outcome). Bartling et al. (2015) show that in sequential decisions pivotal voters are blamed the most for unfair outcomes. Duch et al. (2015) find that in simultaneous collective decisions, proposal power plays an important role for responsibility attribution. In our design with five voters, we can compare the initiator and the pivotal player with a majority voter who is neither. Further, we can distinguish between majority voters who still had a say and those who had no more influence on the outcome. We call the former the intentional voters and and the latter the non-intentional voters.[1] Our experiment shows that punishment targets people who vote for the unfair outcome. People who potentially had an impact on the outcome, and in this respect are intentionally unkind, are punished more, and the pivotal players even more within this group. Analogously, choosing the kind option leads to rewards, which are higher if the choice can be considered as intentional and even higher if the choice is pivotal. We do not find that higher responsibility is assigned to the initiator of a good or bad outcome.

Second, we investigate *individual patterns in responsibility attribution* as all our subjects face several decision situations. Studying different patterns in social interactions has been the object of several studies.[2] We find that the individual behavior can be classified into three main groups. These groups are analogous in the reward and the punishment condition. There is a group of subjects who barely rewards or punishes. Another group of subjects particularly targets the pivotal voter, and a third group of subjects mostly attributes responsibility according to the choices of the voters. In the *Punishment* treatment, we additionally find a small group who focuses on the initiator.

Third, we study *whether responsibility attribution differs for good and bad outcomes in comparable situations.* We do so, by making use of our treatments with reward and punishment options, as well as the combination, in which both reward and punishment are available. As mentioned above, the evaluation of responsibility is consistent between reward and punishment. Subjects reward others for good outcomes very similarly to how they punish others for bad

[1]If the decision is already taken, we cannot infer any intention of the voter. For example, if the outcome has already been decided, a vote for the fair outcome does not mean that the player wanted to be kind.

[2]For example, Falk et al. (2008) investigate different patterns of reward and punishment among their subjects. Most participants express both positive and negative reciprocity, while others only show positive or negative reciprocal fairness preferences. Similarly, Leibbrandt and López Pérez (2011) study heterogeneity in costly reward and punishment. Their results indicate that most subjects follow a mixture of outcome-based and reciprocal preferences. Besides observing different patterns of how subjects sanction, Albrecht et al. (2018) go one step further and examine if and how different behavioral patterns are linked within each subject. In a linear public goods game with decentralized punishment they show that for most subjects the cooperation and the punishment patterns are aligned.

outcomes – both on the aggregate level as well as in the individual analysis.

Fourth, we study *the underlying decision process of responsibility attribution.* The possibility to track eye movements enables us to study the information processes that lead to specific decisions. The fixations on different pieces of information are related to the processing of the inspected information (Just and Carpenter, 1980) and indicate the relative importance of specific information for the decision making process (Duchowski, 2017; Rahal and Fiedler, 2019). For example, a person visually neglecting all choices except the first voter, will probably also attribute responsibility mainly based on this information. However, our results show that the gaze analysis does not confirm the behavioral focus on the pivotal player. If any player is more in the focus, then it is the initiator.

Fifth, we investigate *how the voters respond* to the incentives created by responsibility attribution. We find evidence that voters are (at least partially) aware of how responsibility is attributed. In particular, they are aware that pivotality matters and partially use delegation in order to avoid blame for unkind decisions or seek responsibility for kind decisions in order to gain credit. Further, they have longer response times when they are potentially pivotal, i.e., if their decision can finalize the outcome.

The remainder of this paper is structured as follows: Section 2 explains the experimental design used in this study while Section 3 outlines the different motives used to study responsibility attribution and lists our predictions. Our results are presented in Section 4 and Section 5 concludes.

2 Experimental Design

We build on Bartling et al. (2015) who investigated a sequential voting task with punishment. We add a treatment with reward in order to directly compare responsibility attribution for good and bad outcomes, and we increase the number of voters because it allows investigating more roles in this sequential decision process. In our experiment, we randomly assign the role of voter and recipient to subjects. Five voters and five recipients form a group and keep their roles throughout the whole experiment. The five voters sequentially decide between two allocations in order to distribute 50 points among all ten group members. There are two sets of allocations. In one set, voters can choose between a fair allocation, in which all group members receive 5 points, and an unfair allocation, in which the voters receive 9 points each and the recipients receive 1 point each. In the second set, the voters can choose between a fairer allocation, in which the voters receive 6 points and the recipients receive 4 points, and an unfair allocation, in which the voters receive 8 points each and the recipients receive 2 points each. We chose the two sets in such a way that the alternatives create a similar trade-off, and, thus, the two sets can be treated equally. The position of the voters in the voting sequence is randomly determined. Each voter is informed about the decisions of all previous voters in the sequential process before choosing an allocation. A majority rule is applied, which means that the allocation that is chosen by at least three of the five voters is implemented.

The recipients are informed about the individual voting decisions and thus also about the voting outcome. One randomly determined recipient receives an extra point and has the option to sanction the voters. We vary the sanction option across three treatments: Recipients can only punish voters (*Punishment*), they can only reward voters (*Reward*), and they can reward and punish voters (*Both*). In all treatments, the recipient first has to decide whether to sanction the voters at the cost of the extra point by clicking a button. In the second decision, the recipient can then assign 0 to 7 reward and/or punishment points to each voter individually. Figure 1 illustrates an exemplary decision screen of a selected recipient who decided to sanction the voters.

The payoff of each voter is determined by the resulting voting outcome and the reward or punishment points the voter receives from the recipient. Each recipient gets a payoff according to the chosen allocation. The selected recipient can additionally keep the extra point if she decides not to sanction.

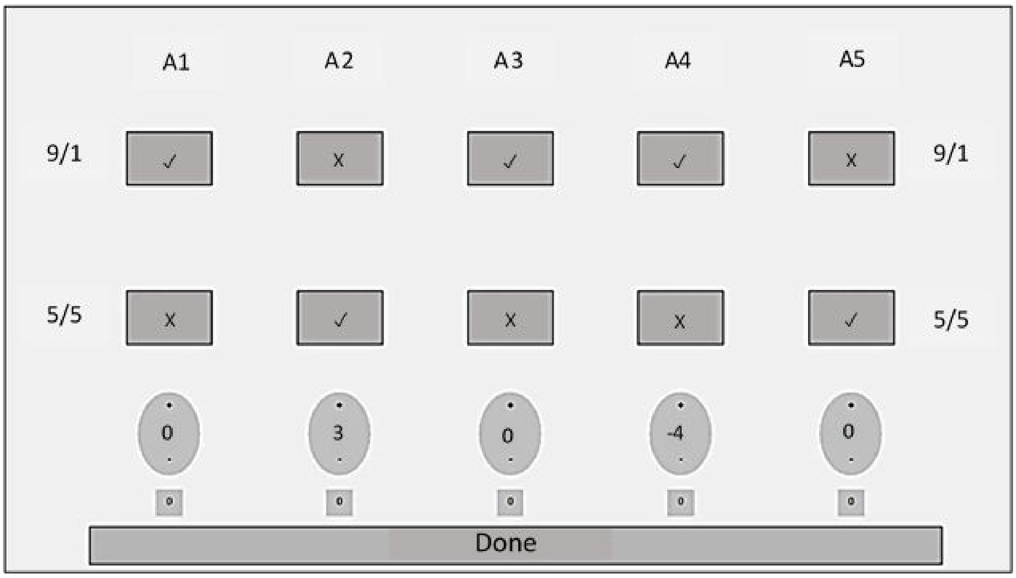


Figure 1: Exemplary decision screen of a recipient in the treatment *Both*.

Note: Voters are denoted as A1 - A5 and their decisions are indicated by a check at the selected allocation. The positioning of the allocations in the top or bottom row was randomly determined. The outcome in this example is the unfair allocation (9 points for each voter and 1 point for each recipient) and the recipient attributes three reward points to voter 2 and deducts four punishment points from voter 4. We added the respective allocation on both sides of the screen to minimize subject's gaze being biased towards one side of the screen. The font size in the figure was enlarged for better readability.

The game is played as a one-shot game and we use the strategy method for both voters and recipients, i.e., each voter and each recipient makes choices for all possible scenarios. Each voter chooses between the fair and unfair allocation in every voter position for every possible combination of previous voter choices. This results in 31 binary choices for each of the two allocation sets which we display in random order. Additionally, the voters play one round of a dictator game for each allocation set resulting in two additional decisions. All recipients act as if they were chosen to be the recipient to sanction. For every possible voting sequence, the recipients decide whether they want to sanction any of the voters and if yes, by how much they want to sanction each voter. The scenarios differ in the decision constellations of the voters and the allocation sets.

As process measures, we collect response time data for both voters and recipients. In addition, we use eye-tracking to record the gaze pattern of the recipients to evaluate the information recipients use when attributing responsibility to the voters.

2.1 Procedural Details

The experiment was programmed using the software "z-Tree" (Fischbacher, 2007). Participants were students who were recruited by the data-system ORSEE (Greiner, 2015). In total, nine sessions were conducted in February 2019, three sessions for each treatment. The experiment was carried out at the experimental laboratory of the University of Konstanz (Lakelab) in Germany. Each session consisted of two groups, 10 voters and 10 recipients, such that there were 30 voters and 30 recipients in each of the three treatments (*Punishment*, *Reward*, *Both*). One subject was excluded from the analysis due to insufficient attention during the session. The average age of our subjects was 22 years (min: 18, max: 33) and 55.56% of the subjects were female. The subjects earned on average 22.58 EUR (about 25.40 USD at that time) which included a show-up fee and an extra compensation for the usage of eye-tracking. The sessions lasted 90 minutes on average.

We used Tobii EyeX eye-trackers with a sampling rate of 60Hz to record gaze data. The subjects used chin rests to improve data quality and the seating distance to the screen was approximately 58 cm. The screens were 22 inch color monitors with a resolution of 1920x1080 pixels. The calibration of the subjects to the eye-tracking system was done at the beginning of the experiment via a seven-point calibration. Two additional subjects have been excluded from the eye-tracking analysis because of technical issues and poor quality of the gaze data. Fixations are identified with the help of the DBSCAN-algorithm (Ester et al., 1996). We create ten non-overlapping areas of interest (AOI), each with a radius of 90 pixels. Each AOI covers a box on the decision screen indicating whether the voter voted in favor or against the specific allocation (see Figure 1). Therefore, for each voter there are two AOIs. The horizontal distance between the centers of two AOIs was 320 pixels and the font size of the cues was set to 20.

3 Criteria and Theoretical Predictions for Responsibility Attribution

In this section, we present criteria according to which people could attribute responsibility. We will assess the relevance of these criteria by using them as predictors for reward and punishment.

Outcome. Outcome-oriented models such as Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) are quite popular. In our voting game, it means that reward and punishment is not directed to a specific voter. The assessment only depends on the outcome. The unfair alternative is considered as bad, and the fair alternative as kind or neutral.[3] The models predict equal punishment for all voters in case of an unfair outcome and if at all, equal reward in case of a fair outcome.

Choice. This motive assumes that a vote for the fair allocation is perceived as kind and a vote for the unfair allocation as unkind, independent of whether the vote was relevant for the outcome or not. It can be considered as a naive notion of intention. If the voter would not take the behavior of the other voters into account and would believe that their vote is decisive, then their vote would correspond to their preference and the vote would express their intention.

Intention. This motive captures preferences as suggested in the reciprocity models of Rabin (1993) and Dufwenberg and Kirchsteiger (2004). Since there are only two options, either both are neutral, or one option is kind and the other is unkind. Theoretically, it depends on the belief of what the other voters do whether voting for fair or unfair is kinder.[4] However, in the experiment, voting for the fair allocation always results in a higher probability of getting the fair outcome than voting for the unfair allocation. Thus, as long as a majority for one of the allocations is not reached, votes are impactful and voting for the fair allocation can be considered as kind and choosing the unfair allocation as unkind. As soon as the decision is made and the outcome can no longer be changed, votes are no longer impactful, no intentions can be inferred, and the vote is considered as neutral.

Initiation. This motive is motivated from Duch et al. (2015) who showed that proposal power is an important aspect in how subjects attribute responsibility for collective decisions. Applied to our experiment, it would assume that the first voter who votes for the resulting outcome has a special responsibility for the outcome.

Pivotality. This motive is motivated from Bartling et al. (2015) who found that the pivotal voter is punished more than the non-pivotal players for the choice of an unfair outcome. The pivotal voter is the third voter who votes in favor of one of the allocations. After this choice, the outcome is determined and can no longer be changed.

Bartling and Fischbacher (2012) Responsibility Measure. In the responsibility measure formalized by Bartling and Fischbacher (2012) (from now on called BF Responsibility), the responsibility of the different voters for a certain outcome is assigned proportional to how much their vote contributes to an increase in the probability that this outcome results. In order to simplify the formulation, we present the case when the outcome is unfair. The measure works exactly in the same way for the fair outcome. The measure depends on the belief about the voter's decisions. It is calculated as follows: First, we calculate for every decision node the probability that an unfair outcome results. Next, each action gets a raw responsibility, which is the difference between the probability before and after the action. Finally, the responsibility measure is the normalized version of the raw responsibility. This means specifically: The responsibility of an action that does not increase the probability of a unfair outcome is set to

[3] According to outcome oriented models, voting for the fair alternative can be considered as neutral and therefore fair outcomes might not be rewarded.

[4] For example, if voters 2 to 5 always vote against voter 1, then the fair vote of voter 1 would actually be unkind.

zero. The responsibility of an action that increases the probability of an unfair outcome is the raw responsibility divided by the sum of all positive raw responsibilities along the decision path from the start to the final outcome. This measure lies between zero and one. As mentioned above, it depends on a belief about the voters' decisions. Practically, we use the empirical distribution of voters' decisions as their belief. Note that the responsibility measure refers to the outcome. We expect that responsibility for the unfair outcome triggers punishment and responsibility for the fair outcome triggers reward.

Engl (2018) Responsibility Measure. Another notion of responsibility has been suggested by Engl (2018). We do not present the full model but explain what it predicts in our case. The idea of the model can be considered as a measure of pivotality. Full pivotality means that the outcome results if the agents performs the action, and the outcome does not result if the alternative action is taken. The measure of Engl (2018) uses the difference between the probability that the outcome occurs if the action is performed with the probability that the outcome occurs if the alternative action is performed. Engl (2018) distinguishes between ex-ante and ex-post causal responsibility. For the calculation of ex-post causal responsibility, we take the outcome as given. This means that the model defines the responsibility of each agent in a voting sequence. An action is pivotal if voting for the other option would have changed the result. Consider a case, in which the unfair outcome resulted. In this case, the probability that the fair outcome would have resulted if a voter voted for the alternative is used as the measure of responsibility of this voter. This definition is very intuitive if one considers the border cases. If a change of the action will lead to the fair outcome for sure, then responsibility equals one. This is, e.g., the case for the third vote for unfair (U) in the sequence FUFUU. If a change of the action will impossibly lead to the fair outcome, then responsibility equals zero. This is, e.g., the case for the last unfair vote (U) in the sequence FUUUU.[5] The responsibility for fair outcomes is defined analogously as the probability that the unfair outcome would have resulted if a voter voted for the alternative. Note that the ex-post responsibility is defined for an action within a decision path. The ex-ante responsibility is calculated as the expected value of the ex-post responsibility of all paths following the actual decision. Thus, the ex-ante responsibility is defined for the action - independent of the path.

Table 1 shows a summary of the presented criteria and their theoretical predictions which we use to analyze how people attribute responsibility. The third column indicates the possible voter positions affected by each criteria.

[5] As a side remark, the first fair vote (F) does not have zero responsibility. Also with the alternative U the fair outcome has some positive probability.

Table 1: Overview of Responsibility Measures

Measure	Theoretical Prediction - Who is responsible?	Voter Positions
Outcome	All voters are responsible for the final outcome.	1,2,3,4,5
Choice	All voters who vote for the final outcome are responsible.	1,2,3,4,5
Intention	Impactful voters are more responsible than non-impactful voters.	1,2,3,4,5
Initiation	The first voter voting for the final outcome is responsible.	1,2,3
Pivotality	The third voter voting for the final outcome is responsible.	3,4,5
BF Responsibility	Voters are more responsible if they increase the probability that the final outcome results.	1,2,3,4,5
Engl Responsibility	Voters are more responsible if choosing the alternative action would have resulted in a different outcome with a higher probability.	1,2,3,4,5

4 Results

Our main research focus is how recipients attribute responsibility for sequential collective decisions, which result in fair and unfair outcomes. To do so, we first test the criteria and theoretical predictions stated in Section 3, followed by an integrative model, a heterogeneity analysis and the analysis of the process data. In the analysis of the voters' behavior, we study voting patterns, strategic voting and delegation, and response times. We do not distinguish between the two possible allocation settings because the results are, as expected, quite similar.[6]

4.1 Sanctioning Behavior

For the analysis, we separate the decisions in whether the outcome was fair or unfair, and classify each voter into the majority group (those who voted for the resulting outcome) or into the minority group (those who voted against the resulting outcome). In each voting sequence there are between three and five majority voters and between zero and two minority voters. Among the majority voters, we distinguish between voters who vote before a majority is reached and who are therefore impactful for the final outcome (first three majority voters) and non-impactful voters (the possible fourth and fifth majority voter). Finally, we separate

[6] In the dictator games, 21 out of 90 subjects chose (5,5) when (9,1) was the alternative and 20 chose (6,4) when (8,2) was the alternative.

the impactful voters (first three majority voters) into the initiator, i.e., first majority voter, the second majority voter and the pivotal voter, who is the third majority voter. The impactful voters are named according to their roles in the sequential decision.

Outcome. Figure 2 shows the average punishment and reward points for fair and unfair outcomes across treatments. Note that in the *Both* treatment, recipients could use the seven points for both punishment and reward. Thus, the *Both* treatment is included in both subfigures.

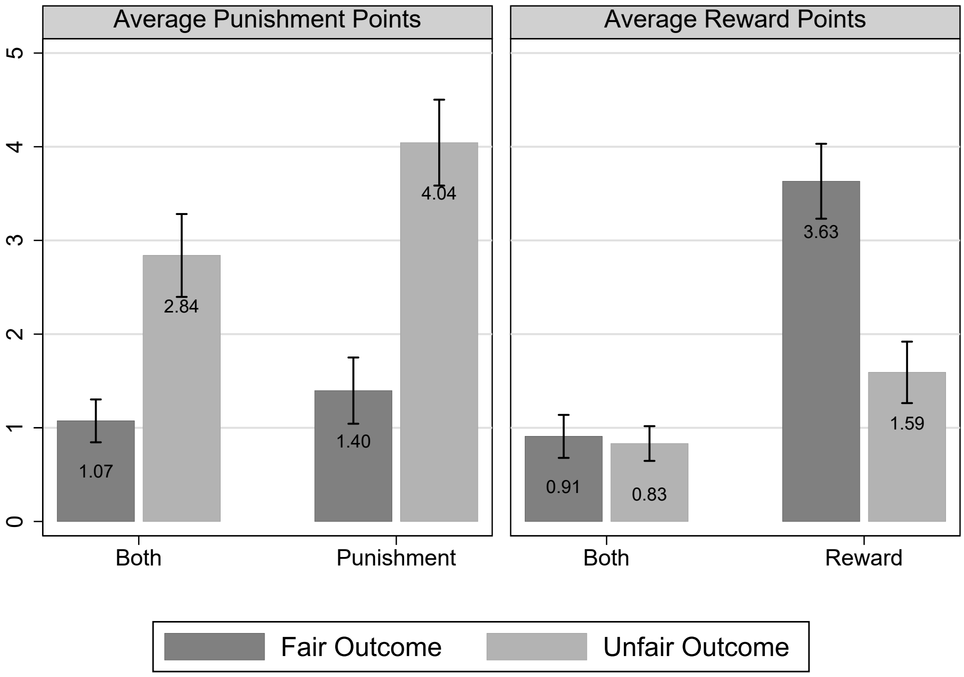


Figure 2: Average punishment and reward points for fair and unfair outcomes across treatments.

Note: Standard error bars are shown in black.

In the *Punishment* treatment, recipients punish more in unfair outcomes than in fair outcomes (4.04 and 1.40 points, Wilcoxon signed-rank test, $p < 0.001$).[7] In the *Reward* treatment, more reward points are assigned on average in fair outcomes compared to unfair outcomes (3.63 and 1.59 points, $p < 0.001$). In the *Both* treatment, where subjects can both punish and reward voters, more punishment points are used on average for unfair outcomes (2.84) than for fair outcomes (1.07, $p < 0.001$). In contrast, recipients do not reward fair and unfair outcomes differently (0.91 and 0.83 points, $p = 0.975$). This implies that punishment is used in a more differentiated way than reward when both options are available.

Our results show that recipients indeed punish unfair outcomes more than fair outcomes. However, our outcome-based prediction cannot be supported since recipients also punish when the outcome is an equal allocation and use reward points for both outcomes.

[7]For every hypothesis test we use a Wilcoxon signed-rank test for matched samples which is based on average decisions per subject.

Choices. Figure 3 shows the average sanction points for different voter roles for unfair and fair outcomes across treatments. Table A3 in the Appendix lists the corresponding average sanction points in more detail by also taking the voter position into account.

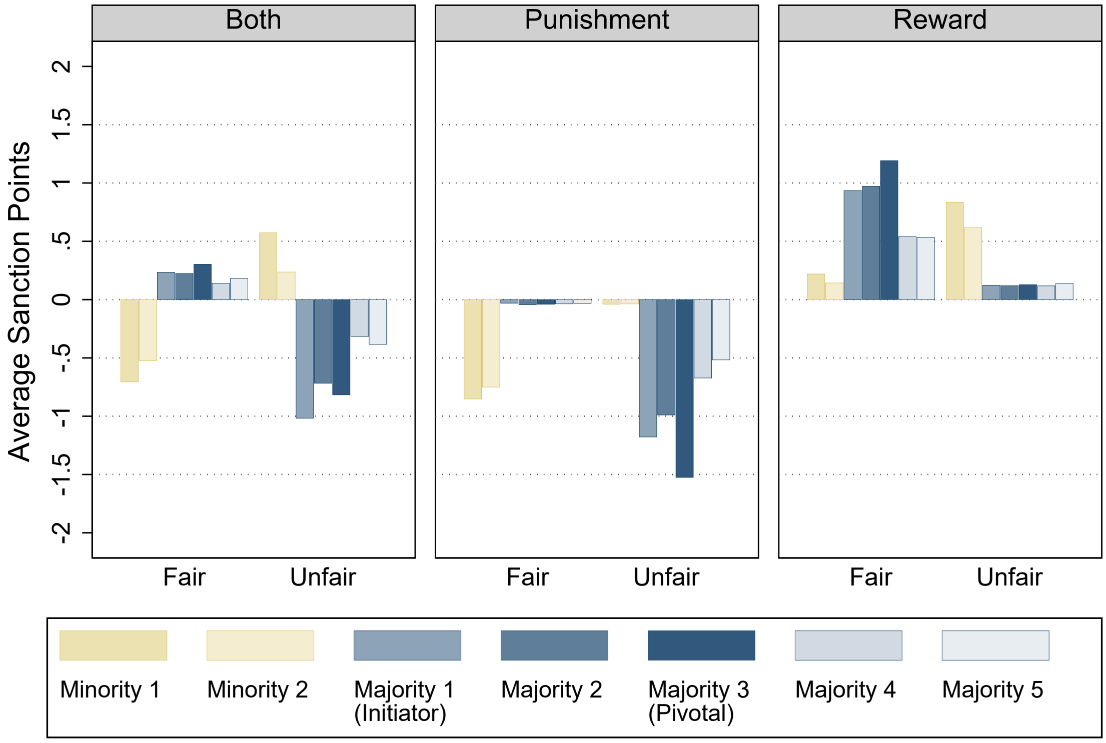


Figure 3: Average sanction points for different voter roles across treatments.

Note: The bars show the average punishment or reward points for different sanction motives separated by outcome (fair vs. unfair) and treatment.

Figure 3 shows that choice clearly matters for responsibility attribution. In all treatments, voting for the fair allocation is rewarded more and/or punished less than voting for the unfair allocation (all *p-values* < 0.02). This is illustrated by comparing the minority voters (yellow bars) with the respective majority voters (blue bars) in Figure 3 for each outcome. Our choice-based prediction can thus be confirmed by the data and shows that subjects attribute responsibility according to choices. However, one should note that punishment points for subjects voting for the fair allocation as well as reward points for subjects voting for the unfair allocation are for almost all voter roles significantly different from zero (all *p-values* < 0.05, except for one case).

Intentions. So far, we have shown that the outcome and the choices of the voters matter. We now disentangle who is held more responsible among the voters choosing the same allocation. We call the first three majority voters of a voting sequence intentional voters, while the majority voters four and five are non-impactful voters as their vote can no longer change the outcome. As shown in Figure 3 recipients punish intentional voters for unfair outcomes and reward them for fair outcomes more than non-impactful voters (all *p-values* < 0.05). In the *Both* treatment, the same results hold except for one case.[8]

Our results show that the intention-based prediction can be confirmed. Thus, when attributing responsibility for an outcome, recipients take the impact of the votes on the final outcome into account.

Initiation. We analyze whether there are differences in perceived responsibility among intentional voters. We first test whether the initiator is sanctioned more than the second majority voter. Across all treatments, we do not find evidence for an initiator effect (average sanction points for initiator vs. second majority voter: *Punishment* -1.18 vs. -0.99; *Reward* 0.94 vs. 0.97; *Both* fair outcomes 0.23 vs. 0.22; *Both* unfair outcomes -1.02 vs. -0.72). Recipients do not seem to punish and/or reward the initiator differently than the second majority voter (all *p-values* > 0.1). Therefore, the initiation-based prediction does not hold.

Pivotality. We expect the pivotal voter to carry the highest responsibility for collective decisions (e.g., Bartling et al. (2015)). We define the pivotal voter as the third majority voter of the voting sequence. In the *Punishment* treatment the pivotal voter is punished more than both other intentional voters when the outcome is unfair (-1.52 vs. -1.18 / -0.99, both *p-values* < 0.02). In the *Reward* treatment the pivotal voter is rewarded the most for fair outcomes (1.19 points on average) which is more than the other two intentional voters (both *p-values* < 0.08). However, in the *Both* treatment the pivotal voter is not treated differently compared to other intentional voters (all *p-values* > 0.1). The pivotality-based prediction can be confirmed partially which suggests that pivotality plays an important role when attributing responsibility for cases where reward and punishment are separately available.

Taken together, our analysis allows us to answer the first research question on how responsibility is attributed to different roles in a decision chain. Our results show that people attribute responsibility differently depending on the outcome of the sequential decision. Generally, subjects are held responsible according to their choices. Furthermore, impactful voters are perceived to be more responsible than non-impactful voters. Last, the pivotal voter bears the highest responsibility while the initiator of a sequential voting sequence is not treated differently than the second impactful voter. In addition, we find that the different criteria of responsibility attribution are the same whether people praise others for good outcomes or blame them for bad outcomes.

4.1.1 Econometric Comparison of Sanctioning Motives

In this section, we provide an econometric comparison between the different motives of how people attribute responsibility. First, we compare the explanatory power of the different motives in isolation in order to study how the criteria help in explaining responsibility attribution.

[8] For fair outcomes the second majority voter and the fifth majority voter are not treated differently.

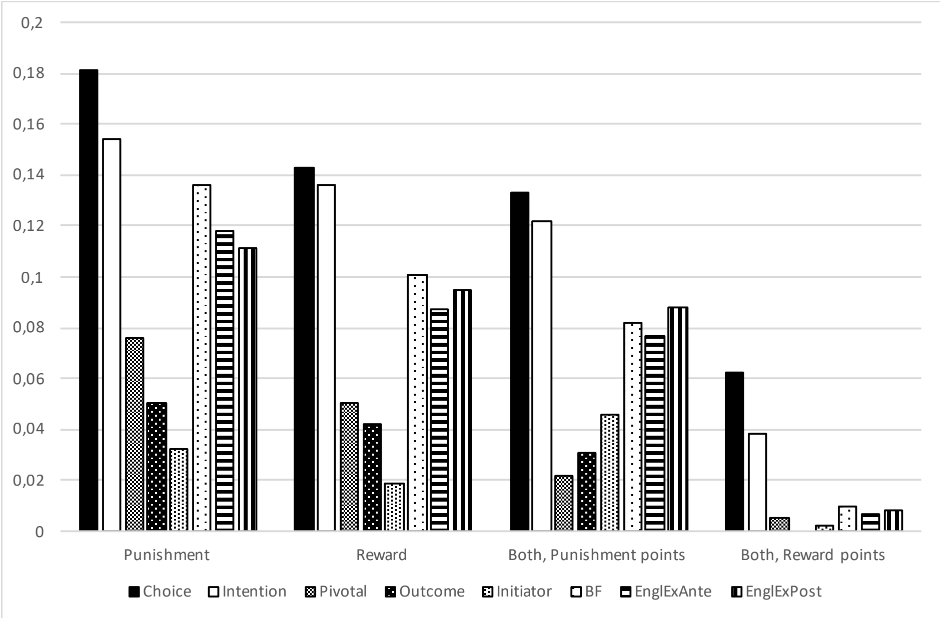


Figure 4: Comparison of R^2 for different responsibility measures

Figure 4 shows the R^2 of individual OLS regressions with reward and/or punishment points as dependent variables and the criteria of responsibility attribution as presented in Section 3 as independent variables.[9] The figure shows that the criterion *choice* has the highest explanatory power across all treatments, followed by the criterion *intention*, further followed by the theories of responsibility attribution (*BF, EnglExAnte, EnglExPost*), which have similar explanatory power. The positional criteria *pivotality* and *initiation* are amongst the criteria with the least explanatory power.

So far, each motive of how responsibility is attributed has been tested separately. Outcome-based models predict when people use punishment or reward, but they do not predict who is held responsible. Models based on reciprocity and intentions can explain who is perceived responsible but not when. We now test which motives have explanatory power when considering all motives simultaneously and, thereby, compare the importance of the different motives. Importantly, even though the outcome is the same in many scenarios in our experiment, the number of votes for and against the outcome differs and also which voter position was associated with which motive (i.e., whether the third, fourth, or fifth voter is pivotal; whether the first, second or third voter is the initiator). These differences might influence how people attribute responsibility in our experiment.

Table 2 shows the corresponding OLS regression outputs. On the left side of the table we regress the punishment points on each sanctioning motive simultaneously for the treatments *Punishment* and *Both*, while on the right side of the table we regress the reward points on each sanctioning motive simultaneously for the treatments *Reward* and *Both*. We control for the size of the majority voters but do not include the voter position.[10]

[9]The output of the individual OLS regressions can be found in the Appendix A.3. Note, that when taken individually, each responsibility measure significantly predicts the attributed sanction points.

[10]We tested whether the voter position influences the perceived responsibility of each voter category by

Table 2: Joint OLS regressions to compare the impact of the criteria on the usage of punishment and reward points

	Punishment Points			Reward Points	
	Punishment	Both		Reward	Both
Choice Unfair	0.797***	0.524***	Choice Fair	0.472***	0.365***
	(0.179)	(0.105)		(0.121)	(0.0878)
Outcome Unfair	-0.0529	-0.0716	Outcome Fair	0.0232	-0.100**
	(0.0709)	(0.0351)		(0.0410)	(0.0322)
Intention Unkind	-0.0617	0.0769	Intention Kind	0.180	0.00851
	(0.139)	(0.0719)		(0.0903)	(0.0404)
Initiator Unfair	0.190	0.262	Initiator Fair	-0.0248	-0.0497
	(0.231)	(0.246)		(0.0625)	(0.0310)
Pivotal Unfair	0.598**	0.183	Pivotal Fair	0.261**	0.0522
	(0.208)	(0.0894)		(0.0807)	(0.0540)
BF Responsibility (U)	0.0332	-0.238	BF Responsibility (F)	0.0926	-0.165
	(0.320)	(0.339)		(0.247)	(0.151)
Ex-ante Engl Resp (U)	-0.0419	-0.156	Ex-ante Engl Resp (F)	-0.565*	-0.299
	(0.222)	(0.311)		(0.238)	(0.192)
Ex-post Engl Resp (U)	0.345*	0.398***	Ex-post Engl Resp (F)	0.425**	0.181
	(0.137)	(0.103)		(0.117)	(0.108)
Size of Majority	-0.000787	-0.0354*	Size of Majority	-0.0125	0.0203
	(0.0240)	(0.0170)		(0.0262)	(0.0123)
Constant	0.0470	0.156*	Constant	0.173	-0.0192
	(0.0958)	(0.0660)		(0.0922)	(0.0579)
Observations	9,600	9,600	Observations	9,280	9,600
R^2	0.210	0.152	R^2	0.165	0.071
Number of Subjects	30	30	Number of Subjects	29	30

Note: OLS fixed effects regressions with punishment points and reward points as dependent variables. Punishment points (left side of the table) can take values from 0 to 7 and are used in the treatments *Punishment* and *Both*. Reward points (right side of the table) can take values from 0 to 7 and are used in the treatments *Reward* and *Both*. *Choice (Un)fair* equals 1 if the (un)fair allocation is chosen. *Outcome (Un)fair* is a dummy that equals 1 if the (un)fair outcome is implemented. *Intention (Un)kind* equals 1 if a voter votes for the (un)fair allocation while no majority was reached before. *Initiator (Un)fair* equals 1 if a voter is the initiator for the (un)fair outcome. *Pivotal (Un)fair* is an indicator that equals 1 if a voter is pivotal for the (un)fair outcome. *BF Responsibility (Un)fair* and *Ex-ante and ex-post Engl Responsibility (Un)fair* correspond to the responsibility measures explained in Section 3. *Size of Majority* indicates the number of majority voters and can take values from 3 to 5.
Robust standard errors in parentheses clustered at the subject level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

regressing the sanction points for each voter role on the voter position. The results show that only for the initiator and the pivotal voter the positioning has an impact on the sanction points. Initiators on position 3 are punished more for bad outcomes than initiators on position 1 and 2. Initiators on position 2 and 3 are rewarded more for good outcomes than initiators on position 1. Pivotal voters on position 5 are punished more than pivotal voters on position 4 in unfair outcomes. Pivotal voters on position 5 are rewarded more than pivotal voters on position 3 and 4 in fair outcomes.

The regression output shows that choices have an explanatory power on top of all other motives. Unfair choices predict the punishment patterns seen in the treatments *Punishment* and *Both*. In both treatments voters who choose the unfair allocation get 0.797 / 0.524 more punishment points than voters choosing the fair allocation. On the other side, fair choices are a good predictor for how people reward collective decisions in the treatments *Reward* and *Both*. In these treatments voters earn 0.472 / 0.365 more reward points when choosing the fair allocation compared to the unfair allocation. In the treatments *Punishment* and *Reward* pivotality has predictive power for the perceived responsibility when considering all motives. Being pivotal for the unfair outcome leads to 0.598 more punishment points and being pivotal for the fair outcome leads to 0.261 more reward points compared to other intentional voters. Looking at the responsibility measures by Bartling and Fischbacher (2012) and Engl (2018) one can see that the criterion *Ex-post Engl* for fair and unfair outcomes helps in explaining the punishment and reward behavior in this joint regression. Subjects with a higher ex-post responsibility are punished and rewarded more for the respective outcomes. The remaining responsibility measures *BF Responsibility* and *Ex-ante Engl* do not help much in explaining who is held responsible when combining all measures. This is because the responsibility measures encompass various individual sanction motives which are included in this regression by other variables.

The results of the econometric comparison indicate that in sequential decisions, subjects mainly focus on the choices and the pivotal decision-maker when attributing responsibility.

4.1.2 Heterogeneity

So far, our analysis was based on average decisions across subjects. We now study the heterogeneity of responsibility attribution by looking at each recipient's individual decisions.

Procedure. In order to detect different sanctioning patterns across subjects we perform a cluster analysis. We first regress the sanction points on the first five sanction motives stated in Section 3 (outcome, choice, intention, initiation, pivotality) for each recipient individually. Second, we use these coefficients in a hierarchical clustering. To do so, we first detect outliers by using the single-linkage method in combination with a euclidean distance measure.[11] We then perform a cluster analysis using the ward-linkage method in combination with a euclidean distance measure. The corresponding dendrogram and the Duda-Hart Je(2)/Je(1) index (Duda et al., 2012) are used to identify the optimal number of clusters.

[11] One recipient in the *Both* treatment and one recipient in the *Reward* treatment were detected as outliers and excluded from the heterogeneity analysis.

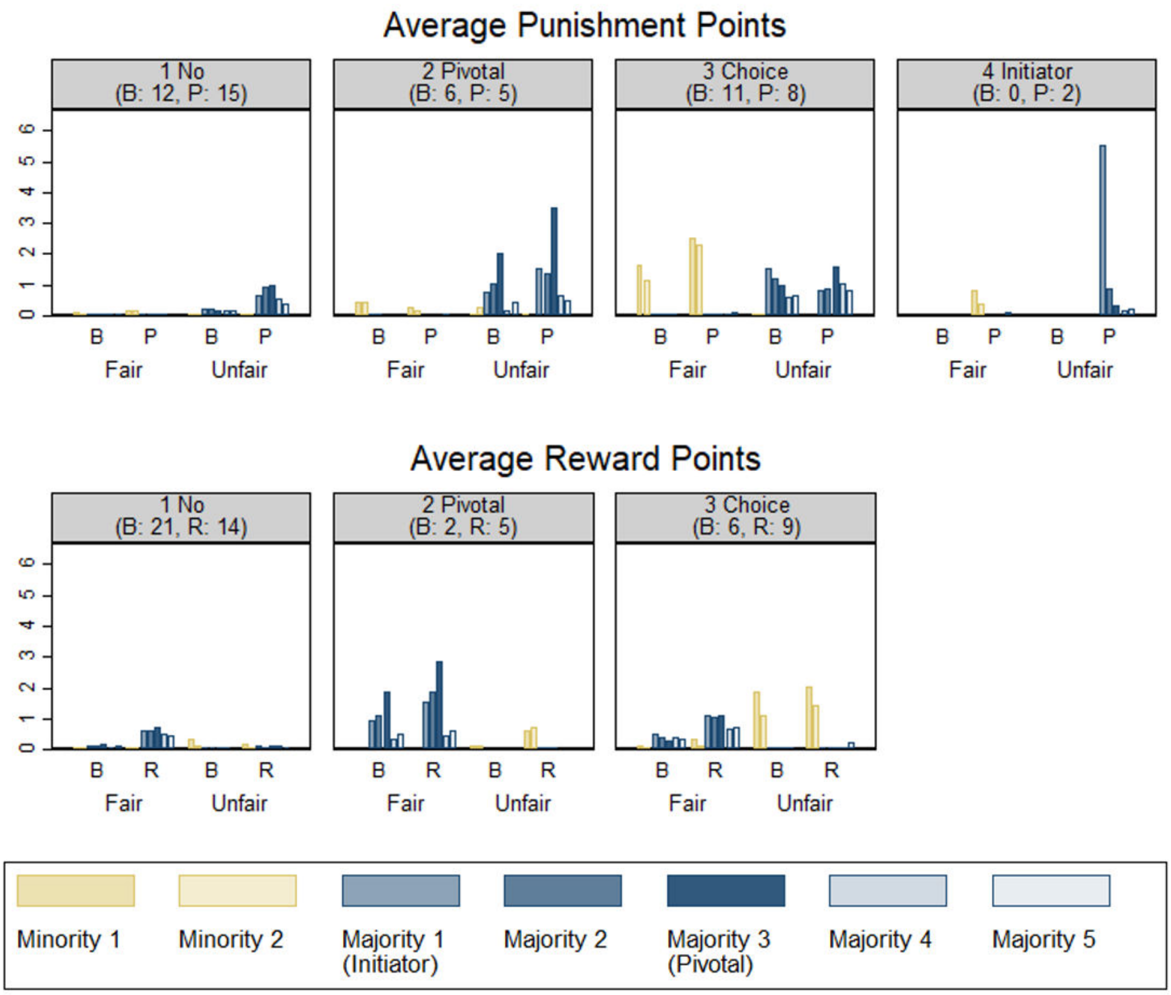


Figure 5: Cluster analysis: Punishment and reward patterns across treatments based on individual sanctioning motives

Note: The figure shows the average punishment and reward points in absolute terms used in each cluster for fair and unfair outcomes across treatments. Hereby, the punishing patterns in the treatments *Punishment* and *Both* are presented in the upper part of the figure, while the reward patterns in the treatments *Reward* and *Both* are presented in the lower part of the figure. The number of subjects contained in each cluster per treatment are indicated in titles of each sub-figure. B = *Both*, P = *Punishment*, R = *Reward*.

Results. The main focus of the heterogeneity analysis is captured by the following three questions: (i) Which punishment and reward types exist in our experiment? (ii) Do the same patterns occur when punishment and reward are used separately compared to simultaneously? (iii) Do subjects punish others according to the same pattern as they reward others? Figure 5 shows the average punishment and reward points for the different voter categories used by recipients categorized within the same cluster. The cluster analysis in the *Both* treatment is performed separately for punishment and reward points and the results are compared to the patterns that are shown in the other two treatments.

(i) Which punishment and reward types exist in our experiment? Four different punishment and three different reward patterns are present in our experiment. Considering only punishment

points, 50% of the subjects in the *Punishment* treatment use no or little punishment (Cluster 1).[12] Subjects categorized in Cluster 2 punish intentional voters and especially the pivotal voter for unfair outcomes. Subjects in Cluster 3 focus on punishing unfair choices, while Cluster 4 consists of two subjects who focus on punishing the initiator. Excluding the two subjects from Cluster 4 does not change the classification of the other subjects.

The cluster analysis in the *Reward* treatment shows that half of the subjects in our experiment only use little reward (Cluster 1). Subjects in Cluster 2 focus on rewarding intentional voters for fair outcomes (especially the pivotal voter), while subjects in Cluster 3 have a preference for rewarding fair choices in both outcomes. Overall, the patterns shown in the treatments *Punishment* and *Reward* are very similar as these clusters can be described by: no or little punishment/reward, focus on the pivotal voter, and focus on choices.

(ii) Do the same patterns occur when punishment and reward are used separately compared to simultaneously? In Figure 5 the respective clusters of the treatments *Punishment* and *Reward* are displayed next to the patterns of the *Both* treatment. The punishing behavior of recipients in the *Both* treatment can be described by the same patterns as in the *Punishment* treatment (except for the fourth cluster which does not exist in the *Both* treatment). Similarly, the three rewarding patterns identified in the *Both* treatment match the rewarding patterns in the *Reward* treatment.[13] Therefore, we identify very similar punishment and reward patterns irrespective of whether the options are separately or simultaneously available.

(iii) Do subjects punish others according to the same pattern as they reward others? To put it in other words: does a subject who punishes mostly the pivotal voter for unfair outcomes also reward mostly the pivotal voter for fair outcomes? With the help of the *Both* treatment we can answer this question by looking at the within-subject comparison of expressed punishment and reward pattern. Table 3 shows the distribution of recipients in the *Both* treatment into the identified punishment and reward patterns. Overall, 15 subjects follow the same pattern when punishing and rewarding. The remaining 14 subjects use different patterns for punishment and reward. To test whether the amount of subjects using matching patterns is significant we use a variation of the Fisher's exact test for our 3 x 3 contingency table with fixed border probabilities. Since we know which patterns correspond to each other, we include this information and compute the likelihood of getting more matching subjects than observed in our experiment. The probability of observing at least 15 subjects who use the same reward and punishment pattern given the frequencies we observe for each pattern individually is 0.076. The significance is not stronger due to subjects who punish unfair choices but do only use little reward. However,

[12] There are two subjects who do not use any punishment point across all decisions in the *Punishment* treatment.

[13] Note that the magnitudes of average punishment and reward points shown in Figure 5 for each cluster in the *Both* treatment are smaller than the corresponding comparison in the *Punishment* and *Reward* treatments. This can be explained by our design since in all treatments the maximum number of sanction points is limited to seven. This means that recipients in the *Both* treatment can use up to seven point for rewarding and punishing while in the other two treatments the seven points can be used for the single sanction option. Also note that while on an aggregate level the average reward points used in the *Both* treatment are not similar to those in the *Reward* treatment (see Figure 2) the individual reward patterns are very similar. This is due to the different number of subjects contained in each cluster for the treatments *Both* and *Reward*.

we observe a high number of subjects who punish and reward in the same way suggesting that subjects often align their rewarding and punishing pattern.

Table 3: Distribution of reward and punishment patterns in the treatment *Both*

		Reward Patterns			
		Little Reward	Pivotal	Choices	Total
Punishment Patterns	Little Punishment	10	0	2	12
	Pivotal	3	2	1	6
	Choices	8	0	3	11
	Total	21	2	6	29

In sum, the presented heterogeneity analysis sheds light on our second and third research focus on disentangling different types of responsibility attribution and whether the process of responsibility attribution differs for blaming and praising others. We identify very similar punishment and reward patterns in our experiment. The expressed patterns are: little sanctioning, sanctioning according to intentions / pivotality and sanctioning according to choices. This result is in line with the average behavior across all subjects shown in the regression Table 2. The patterns for punishing and rewarding in the *Both* treatment are also very similar to the patterns found in the respective treatments. Further, over 50 % of recipients follow the same pattern when punishing and rewarding. Although we observe a high frequency of aligned sanctioning patterns, there are still recipients who follow different patterns when punishing and rewarding.

As a robustness and validation check for the responsibility measures we perform a second cluster analysis only based on the variables *BF Responsibility* and *Engl Responsibility.* These results can be found in the Appendix A.4. The responsibility measures manage to identify the same behavioral patterns and most subjects are categorized into the same clusters. Therefore, the responsibility measures capture similar behavioral patterns across subjects as the individual sanctioning motives and represent a combination of the individual sanctioning variables.

After having analyzed how recipients on an aggregate and individual level attribute responsibility we complement the behavioral analysis by looking at the behavior of the voters.

4.2 Voting Behavior

We now turn to the voting behavior and analyze how subjects vote in collective decisions under the prospect of reward and punishment. Pivotality is an important aspect in the process of responsibility attribution as we have shown in the analysis of the recipients' behavior. Accordingly, strategic voters might take this into account and prevent (favor) being the target of punishment (reward) linked to pivotality.

We examine this strategic behavior by studying the voters' behavior in potentially pivotal decisions.[14] Table 4 offers a general overview of the share of decisions in which voters choose the unfair allocation depending on their position and the previous votes. We separate the results by our treatments (columns 3-5) and by the decisions of the voters in the two dictator decisions (columns 6-8). Importantly, we use the dictator game to elicit the preference of the voters for the fair or unfair allocation when neither a collective decision nor punishment or reward are implemented. 63 voters show a preference for the unfair allocation, while 14 voters show a fair preference. The remaining 13 voters have mixed preferences depending on the two allocation sets we offer them. Bold sequences in column 2 indicate situations in which voters face a potentially pivotal decision.

Table 4: Voting Behavior - Share of unfair choices

		By treatment			By true preference		
Voter	Previous Voters	Both	Punishment	Reward	Fair (14 Voters)	Mixed (13 Voters)	Unfair (63 Voters)
1	-	0.45	0.53	0.43	0.07	0.58	0.54
2	U	0.55	0.63	0.40	0.00	0.69	0.61
	F	0.48	0.47	0.42	0.04	0.54	0.53
3	**UU**	0.40	0.50	0.48	0.11	0.62	0.51
	FF	0.35	0.28	0.40	0.07	0.46	0.38
	Tie	0.48	0.60	0.49	0.11	0.60	0.60
4	UUU	0.23	0.35	0.40	0.18	0.54	0.32
	FFF	0.02	0.02	0.12	0.00	0.00	0.07
	2 of 3 U	0.42	0.52	0.53	0.07	0.46	0.59
	2 of 3 F	0.47	0.44	0.52	0.13	0.44	0.56
5	3 or more U	0.31	0.30	0.40	0.26	0.52	0.31
	3 or more F	0.05	0.03	0.14	0.02	0.05	0.09
	Tie	0.48	0.54	0.61	0.11	0.47	0.65

Note: Bold marked sequences in column 2 indicate decisions in which voters are potentially pivotal.

Voters who face a potentially pivotal decision are influenced by the choices of the previous voters. Across all treatments, the share of unfair choices for potentially pivotal voters is higher when the majority of previous voters voted unfair in comparison to a fair majority of previous votes (columns 3-5).[15] The biggest discrepancy results in the treatment *Punishment* for potentially pivotal voters on voting position three. 50 % of these voters' decisions are unfair when the first two voters voted unfair, while only 28 % of the decisions are unfair when the first two

[14]Potentially pivotal means that exactly two of the previous voters voted for the same allocation and that the own vote can be deterministic for the outcome. These situations can only appear for voters on positions three, four and five.

[15]One exception are the decisions of potentially pivotal voters in fourth position in the treatment *Both* where 47% of decisions are unfair following a fair majority, while only 42% are unfair following an unfair majority.

voters voted fair.

The sequential decision design allows voters to use strategic (non-)delegation in order to avoid (seek) pivotality. Here, we focus on the potentially pivotal voters on either position three or four. In these cases the voters can ensure being pivotal by following the majority of previous votes. But the voters can also vote against the majority of previous voters and can therefore delegate the notion of being pivotal to the next voter. Subjects showing a preference for the fair allocation in the dictator game mostly choose the fair allocation when being potentially pivotal (column 6 in Table 4). In contrast, voters expressing a preference for the unfair allocation in the dictator game often behave against their true preference in potentially pivotal situations of the collective decision (column 8). On voting positions three and four, potentially pivotal voters with an unfair preference ensure the unfair outcome by being pivotal in only 51% and 59% of all decisions respectively. In 62% and 44% of the cases where the majority of previous voters chose the fair allocation, potentially pivotal voters on positions three and four voted against their true preference and decided on being pivotal for the fair outcome. On voting position 5 where no strategic (non-)delegation is possible voters followed their true preference in only 65% of the cases.

Taken together, voters showing a preference for the unfair outcome often vote against their preference when being potentially pivotal. They avoid (seek) being pivotal for unfair (fair) outcomes by strategic (non-)delegation. But surprisingly many subject do not try to appear fair when the outcome can no longer be changed, in particular when the outcome of the decision is unfair.

This completes our behavioral results. We now turn to the processing data we collected for recipients and voters to complement the behavioral analyses.

4.3 Eye–Tracking Results

We track the gaze of participants in order to analyze whether their sanctioning behavior is reflected in their information search. We collected the data for 90 participants. Three subjects had to be excluded due to poor gaze data quality. The recipients made choices on two decision screens. On the first screen, the recipients made the decision whether to sanction or not. In addition to the votes being displayed, the decision screen had two large bars at the bottom indicating the willingness to sanction or not (see Figure A2 in Appendix A.1). On the second screen, the recipients decided on how many sanction points they wanted to allocate to the voters.[16] The second decision screen showed the votes as well as additional buttons for allocating punishment and reward points (see Figure 1). We focus on the results for the first decision screen and we mention the corresponding results for the second screen in the Appendix A.5. Since on the second decision screen, subjects could allot the sanction points to each voter, it involved a lot of clicking and focus on the buttons.

To examine whether the gaze data of the recipients is in line with their sanctioning behavior, we use the average number of fixations and the dwell time. These measures provide an insight

[16] Only the recipients who decided to sanction moved on to the second screen.

into the attention, the information search process of the participants, and how they interpret the content of scenarios (Rahal and Fiedler, 2019). Table 5 shows the summary statistics for multiple gaze measures per decision separated by treatments and outcomes.

Table 5: Summary Statistics - Eye – Tracking Data Per Decision

	Both		Punishment		Reward	
	Fair	Unfair	Fair	Unfair	Fair	Unfair
Avg. no. of total fixations	22.02	22.74	20.71	22.93	20.95	20.32
SD	16.20	17.12	14.26	16.14	16.55	15.97
N	928	928	960	960	896	896
Avg. no. of AOI fixations	11.06	11.57	10.95	12.14	10.79	10.41
SD	7.92	8.49	7.58	8.47	8.18	8.02
N	893	892	928	932	868	871
Avg. dwell time (ms)	4124.57	4355.59	3950.31	4365.17	4003.98	3813.3
SD	3622.21	3907.97	3341.36	3517.09	3690.84	3582.63
N	928	926	957	957	896	896
Avg. dwell time on AOIs (ms)	1984.84	2123.83	2010.77	2229.70	2033.61	1928.15
SD	1727.44	1913.07	1796.42	1843.51	1898.77	1825.90
N	893	891	928	932	868	871
Avg. dwell time per AOI fixation (ms)	174.77	175.34	175.82	179.76	178.62	175.58
SD	52.23	49.41	48.19	68.50	55.22	51.25
N	893	891	928	932	868	871
Avg. no. of Voters looked at	4	3.98	4.11	4.21	3.97	3.94

Note: All measures display the average per decision. Dwell time is the total time spent on all fixations. Dwell time on AOIs is the total time spent on fixations within the 10 AOIs containing the voting information.

As shown in Table 5, in case of an unfair outcome the average number of total fixations in the *Punishment* treatment is higher than in the *Reward* treatment. The average number of fixations in our AOIs showed a similar trend with voters in the *Punishment* treatment receiving the most fixations when an unfair outcome results. For the *Reward* treatment, the voters receive more fixations in case of a fair outcome as compared to an unfair outcome. This seems plausible as in the *Reward* treatment recipients might focus more on the voters in case of a fair outcome in order to decide whom to reward. These results are in line with our behavioral result depicted in Figure 2, which shows that recipients allocate more reward (punishment) points when the outcome is fair (unfair). We also analyze the average dwell time and average dwell time on our AOIs per decision. The average dwell time on AOIs is longest for an unfair outcome in the *Punishment* treatment. Also, across all our treatments and scenarios, on average, four out of five voters are looked at on the first decision screen. This indicates that the participants were attentive and mostly looked at the whole screen.

It is important to note that certain voter positions can attract relatively more fixations merely due to being at the center of the screen. Also, scenarios in which a voter is salient might attract more attention. For instance, in situations where one voter votes differently than the other four voters, saliency is strong and might influence the gaze pattern of the recipients. In order to understand the importance of positioning and saliency, we conduct a fixed effects regression with the share of fixations as a dependent variable fixing on Voter Position 1. We control for voter position and saliency in this regression and the results are shown in Table 6.

Table 6: Share of Fixations - Importance of Position and Saliency

	Punishment	Reward	Both
Voter Position 2	0.0359**	0.0468**	0.0241
	(0.0122)	(0.0144)	(0.0168)
Voter Position 3	0.130***	0.141***	0.123***
	(0.0229)	(0.0311)	(0.0288)
Voter Position 4	-0.00267	-0.00251	-0.0226
	(0.0145)	(0.0137)	(0.0170)
Voter Position 5	-0.0392*	-0.0544***	-0.0829***
	(0.0179)	(0.0140)	(0.0173)
Only One Salient	0.0526***	0.0684***	0.0714***
	(0.00878)	(0.0119)	(0.00927)
Neighbor of Only One Salient	0.00541	0.00881	-0.00299
	(0.00356)	(0.00534)	(0.00436)
Minority	0.00989*	0.0174*	0.00801
	(0.00410)	(0.00644)	(0.00446)
Constant	0.168***	0.163***	0.185***
	(0.0119)	(0.0128)	(0.0151)
Observations	9,295	8,695	8,925
R-squared	0.151	0.169	0.179
Number of Subjects	30	28	29

Note: Fixed effects regression with share of fixations as dependent variable fixing on Voter Position 1. *Voter Position 2-5* indicate the position of the voter in the decision scenario. *Only One Salient* is a dummy variable indicating if the voter is the only salient voter. *Neighbor of Only One Salient* is a dummy variable that takes the value 1 if the voter is next to the only salient voter. *Minority* is a dummy variable indicating if the voter is part of the minority.
Robust standard errors in parentheses clustered at the subject level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Compared to the share of fixations on Voter 1, voters on position 2 and 3 receive a higher share of fixations while voters on position 5 receive a lower share of fixations. This can be due to the natural reading habit of reading from left to right in Western culture. Also, voters who are part of the minority group receive more fixations due to their saliency. This effect is even stronger when there is only one minority voter. The regression output shows that both saliency and voter position highly influence the gaze data.

Besides the influence of the positional measures of voting sequence and saliency, we will now discuss how the individual sanction motives (discussed in Section 3) impact the fixations. We regress the share of fixations of the recipients on the different individual sanction motives incorporated in each decision screen. In addition, we control for the position and saliency measures mentioned above. The regression output is shown in Table 7 and suggests that voters are focused more when being the initiator. The pivotal voter does not receive a special focus. The result is against our expectations, as the pivotal voter is punished and rewarded the most.

Table 7: Share of Fixations - Impact of Sanction Motives

	Punishment	Both		Reward	Both
Choice Unfair	-0.000884	0.00475	Choice Fair	0.0104	-0.00519
	(0.00628)	(0.00483)		(0.00622)	(0.00455)
Intention Unkind	0.00415	-0.000289	Intention Kind	-0.00694	-0.00710
	(0.00669)	(0.00608)		(0.00742)	(0.00655)
Initiator Unfair	0.0172*	0.0211*	Initiator Fair	0.0244**	0.0158*
	(0.00671)	(0.00766)		(0.00866)	(0.00644)
Pivotal Unfair	-0.00530	-0.00463	Pivotal Fair	0.00563	0.000468
	(0.00663)	(0.00641)		(0.0109)	(0.00909)
Controls	Yes	Yes	Controls	Yes	Yes
Constant	0.160***	0.175***	Constant	0.152***	0.185***
	(0.0114)	(0.0147)		(0.0120)	(0.0155)
Observations	9,295	8,925	Observations	8,695	8,925
R-squared	0.152	0.181	R-squared	0.171	0.180
Number of Subjects	30	29	Number of Subjects	28	29

Note: OLS fixed effects regression with share of fixation of the first screen as the dependent variable fixing on Voter Position 1. The controls include variables for *voter position 2-5, saliency and minority*.
Robust standard errors in parentheses clustered at the subject level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

As the recipients allocate sanction points to individual voters on the second decision screen, it is possible that they might fixate more on the different voters while sanctioning. However, on the second screen, the measures for share of fixations and dwell time show a similar trend as on the first screen and the respective tables can be found in the Appendix A.5. [17]

To summarize, we answer our fourth research question on studying the underlying decision process of responsibility attribution with the help of our analysis of the gaze data. The gaze data show that saliency and positioning are highly important. When controlling for these factors, subjects focus more on the initiator, while the pivotal voters does not receive a special attention.

To complete our results, we now turn to the processing analysis of the voters.

[17]Voter position and saliency strongly influence the gaze pattern of the recipients even on the second decision screen. With respect to the sanction motives, choosing an unfair allocation as well as initiating an unfair outcome leads to receiving a higher share of fixations as compared to the other sanction motives in the *Punishment* and *Both* treatments. In the *Reward* treatment, initiating a fair outcome and intentionally voting for a fair outcome attracts a higher share of fixations. Overall, the gaze pattern on the first and second screen exhibited similar trends.

4.4 Processes in Voting Behavior

For the voters, we collected response time as a process measure. We hypothesize that response times inform us about the decision-making process and the difficulty of the decision (Konovalov and Ruff, 2022; Konovalov and Krajbich, 2019; Hausfeld et al., 2020). A decision has to be considered as difficult if the voter strategically votes against the outcome preferences when being potentially pivotal. These decisions are characterized by a higher internal conflict, and should be accompanied by longer response times (Rubinstein, 2007). In Section 4.2, we showed that voters often decide against their true preference when being potentially pivotal. Potentially pivotal means that exactly two previous voters voted for the same allocation and that the outcome can now be determined by the respective voter. This behavior suggests an internal conflict of being potentially pivotal.

Table 8 shows how the response time of voters is affected by the voter position and being potentially pivotal. When accounting for the position the voter is in, we find that voters take significantly more time in choosing an allocation when they are potentially pivotal.

Together with the results presented in Section 4.2, we can answer our last research question on how voters respond to the incentives created by responsibility attribution. Voters are aware of the responsibility that is linked to pivotality as they strategically use delegation to avoid punishment or non-delegation to gain reward even if it means that they vote against their true preference. This behavior is accompanied by a higher response time.

Table 8: Response Time Analyses Voters

	Dependent Variable: Decision Time Voters			
	Punishment	Reward	Both	Total
Voter Position	0.0451**	0.0137	0.0452***	0.0347***
	(0.0145)	(0.0135)	(0.0109)	(0.00762)
Potentially Pivotal	0.142***	0.0943**	0.0939**	0.110***
	(0.0299)	(0.0330)	(0.0339)	(0.0186)
Constant	1.284***	1.338***	1.244***	1.289***
	(0.0552)	(0.0547)	(0.0421)	(0.0300)
Observations	1,860	1,860	1,860	5,580
R-squared	0.035	0.011	0.021	0.021
Number of Subjects	30	30	30	90

Note: OLS fixed effects regression with Decision Time of Voters (taken in log) as dependent variable. *Voter Position* indicates the position the voter holds in the decision she faces and can range from 1 to 5. *Potentially Pivotal* is a dummy variable indicating whether the voter faces a decision where she can be pivotal depending on her choice.
Robust standard errors are clustered on individuals in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

5 Conclusion

In our study, we use blame and praise to investigate how people attribute responsibility in decision chains. In our experiment, five voters choose sequentially between two options of how to allocate points between voters and recipients. One option is fair, the other is unfair. The recipients can reward and/or punish the voters, which we take as our measure of responsibility attribution. Our results are in line with Bartling et al. (2015) who find that the pivotal voter is assigned more punishment if the voters vote for an unfair allocation. We extend this result, showing that pivotality also matters when the voters vote for the fair allocation and the recipients can reward. This means that pivotality generally matters for the attribution of responsibility. Behind this general result, there is strong heterogeneity in how recipients assign reward and punishment. There is a group who barely rewards or punishes. The second group most strongly sanctions the pivotal voter and to a lesser degree the other voters who voted for the resulting outcome, but barely sanctioned the minority voters. The third group of participants focuses on the choice. This group also punishes the voters of the unfair allocation if the fair allocation results and rewards the voters of the fair allocation if the unfair allocation results. In the *Punishment* treatment there was a fourth group who mainly punished the initiator of the unfair allocation. Even though people have rather sophisticated responsibility attribution patterns, the conceptual models of Bartling and Fischbacher (2012) and Engl (2018) explained surprisingly little in comparison to the simple and mechanistic idea that it is the deed that determines responsibility. Our results also show that the voters are aware of how responsibility is attributed and are particularly concerned when they are pivotal, which is also visible in longer response times.

Of course, specific features of the experiment could be relevant for the outcome, for example when investigating the role of the initiator. In our experiment, the two options were chosen with similar frequency. It is possible that when an action is rarely chosen, the initiator is assigned a higher level of responsibility. This could in particular be relevant in the case of fatalities. Nevertheless, using the strategy method as in our experiment would allow experimentally studying such an environment, even though the interesting cases occur rarely. Another feature of our design could have reduced the relevance of the pivotal voter. Voters still had to vote, even when the decision was already made. Finishing the procedure when the result is determined could make the pivotal voter even more focal. Think of penalty shootouts in soccer where the pivotal player is usually the one who is praised or blamed most. Such variants could reveal the sensitivity of responsibility attribution to the specific situation. In addition, investigating costless reward and punishment could reveal whether selfish people apply different patterns of responsibility attribution. However, one can argue that also outside of the lab, reward and punishment bear some cost and therefore it is more important to know the responsibility attribution of people who are willing to bear such cost.

The responsibility attribution in our voting game captures situations, in which people take sequential decisions or actions that jointly generate an outcome. What does this imply in the real world, for example in the case of a disaster? First, any bad action is attributed some responsibility. Second, the pivotal person, i.e., the person after who's action the disaster was unavoidable is generally assigned the highest responsibility. In his book, Whittingham (2004) observes that often the institutional environment is an important reason for disasters. Translated to our setting, he would consider the initiator as particularly responsible. However, we find few subjects who agree with this view. Of course, there are important differences to our lab experiments. First, in such disasters it is more difficult to identify the sequence. In particular, it is not easy to identify the pivotal agent - it is difficult to find out when the disaster was no longer avoidable. Further, the different agents are less symmetric, both with respect to their contribution to the disaster and with respect to their formal responsibility. There are also many institutional details that could matter. Further research will allow investigating such variations, and our experiment provides a framework to do so.

References

Albrecht, F., Kube, S., and Traxler, C. (2018). Cooperation and norm enforcement - The individual-level perspective. *Journal of Public Economics*, 165:1–16.

Bartling, B. and Fischbacher, U. (2012). Shifting the blame: On delegation and responsibility. *Review of Economic Studies*, 79(1):67–87.

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

Besley, T. (2006). *Principled Agents?: The Political Economy of Good Government*. Oxford University Press on Demand.

Bolton, G. E. and Ockenfels, A. (2000). ERC: A Theory of Equity, Reciprocity, and Competition. *The American Economic Review*, 90(1):166–193.

Charness, G. (2000). Responsibility and effort in an experimental labor market. *Journal of Economic Behavior & Organization*, 42(3):375–384.

Duch, R., Przepiorka, W., and Stevenson, R. (2015). Responsibility Attribution for Collective Decision Makers. *American Journal of Political Science*, 59(2):372–389.

Duchowski, A. T. (2017). *Eye Tracking Methodology: Theory and Practice: Third Edition*. London: Springer-Verlag Ltd.

Duda, R. O., Hart, P. E., and Stork, D. G. (2012). *Pattern Classification*. John Wiley & Sons.

Dufwenberg, M. and Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior*, 47(2):268–298.

Engl, F. (2018). A Theory of Causal Responsibility Attribution. SSRN Scholarly Paper ID 2932769, Social Science Research Network, Rochester, NY.

Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In Simoudis, E., Han, J., and Fayyad, U., editors, *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, pages 226–231. AAAI Press.

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

Fehr, E. and Schmidt, K. M. (1999). A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics*, 114(3):817–868.

Feinberg, J. (1970). *Doing & Deserving; Essays in the Theory of Responsibility*. Princeton University Press.

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

Gerstenberg, T., Lagnado, D. A., Speekenbrink, M., and Cheung, C. (2011). Rational order effects in responsibility attributions. *Proceedings of the 33rd Annual Conference of the Cognitive Science Society.*, 33(33):1715–1720.

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

Hart, H. L. A. and Gardner, J. (2008). *Punishment and Responsibility: Essays in the Philosophy of Law*. Oxford University Press, Oxford.

Hausfeld, J., Fischbacher, U., and Knoch, D. (2020). The value of decision-making power in social decisions. *Journal of Economic Behavior & Organization*, 177:898–912.

Iyengar, S. (1994). *Is Anyone Responsible?: How Television Frames Political Issues*. University of Chicago Press.

Just, M. A. and Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4):329–354.

Konovalov, A. and Krajbich, I. (2019). Revealed strength of preference: Inference from response times. *Judgment & Decision Making*, 14(4).

Konovalov, A. and Ruff, C. C. (2022). Enhancing models of social and strategic decision making with process tracing and neural data. *Wiley Interdisciplinary Reviews: Cognitive Science*, 13(1):Epub ahead of print.

Leibbrandt, A. and López Pérez, R. (2011). Individual heterogeneity in punishment and reward. Technical report, UAM. Departamento de Análisis Económico, Teoría Económica e Historia Económica.

Rabin, M. (1993). Incorporating Fairness into Game Theory and Economics. *The American Economic Review*, 83(5):1281–1302.

Rahal, R.-M. and Fiedler, S. (2019). Understanding cognitive and affective mechanisms in social psychology through eye-tracking. *Journal of Experimental Social Psychology*, 85:103842.

Ross, L. and Nisbett, R. E. (1991). *The Person and the Situation : Perspectives of Social Psychology*. New York (N.Y.) : McGraw-Hill.

Rubinstein, A. (2007). Instinctive and cognitive reasoning: A study of response times. *Economic Journal*, 117(523):1243–1259.

Weiner, B. (1995). *Judgments of Responsibility: A Foundation for a Theory of Social Conduct*. New York: Guilford Press.

Whittingham, R. (2004). Design Errors. In *The Blame Machine: Why Human Error Causes Accidents*. Routledge.

A Appendix

A.1 Supplemental Material: Experimental Design

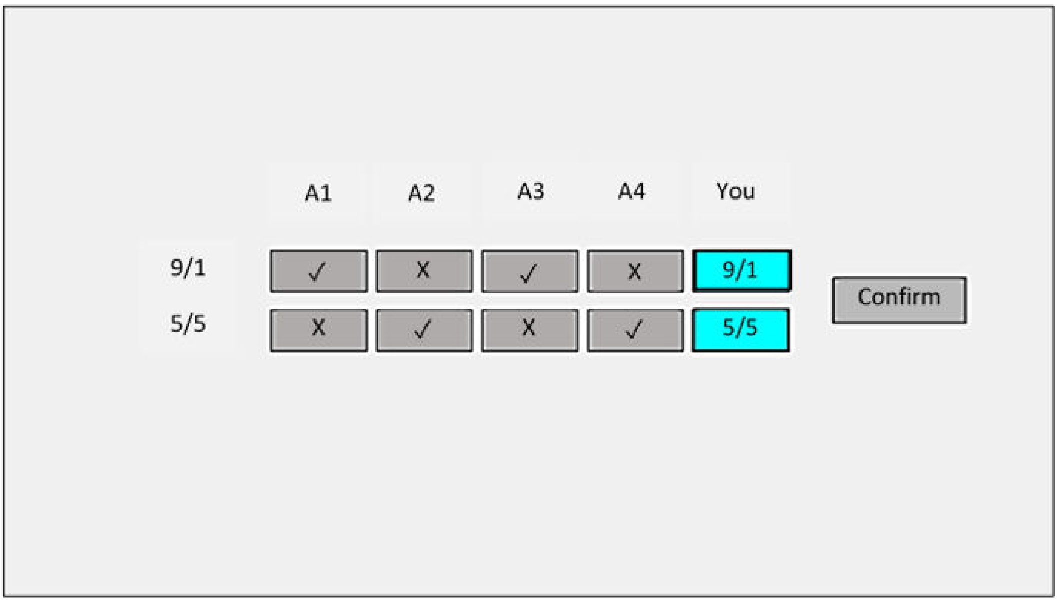


Figure A1: Exemplary screen for a voter

Note: The voter is on position five. Original text translated into English and font size enlarged for better readability.

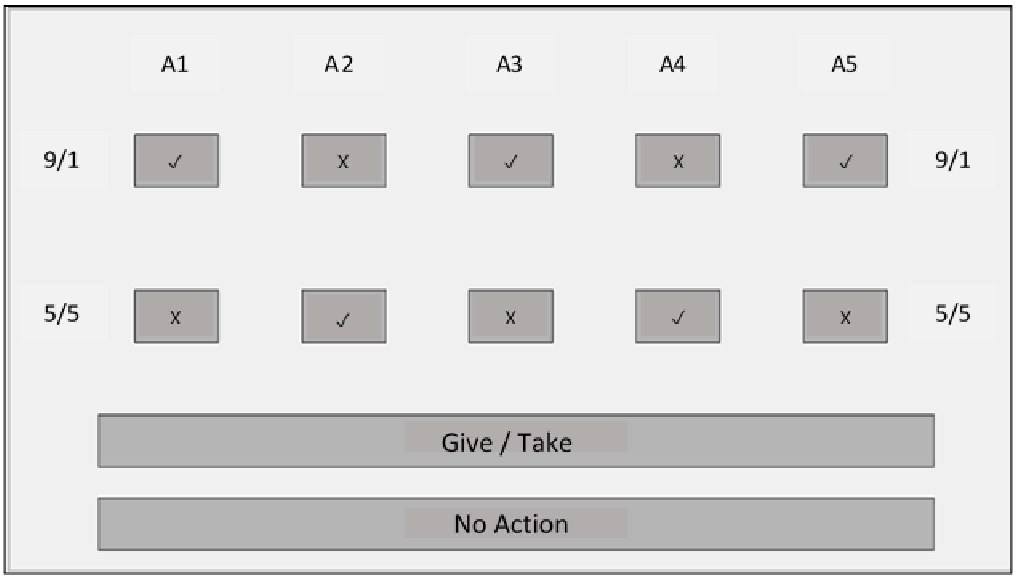


Figure A2: Exemplary first decision screen of a recipient

Note: The recipient is in the *Both* treatment. Original text translated into English and font size enlarged for better readability.

Table A1: Scenarios of Voters

Voter[1]	Voter[2]	Voter[3]	Voter[4]	Voter[5]	Scenario	Voter Position
-	-	-	-	-	1	1
2	-	-	-	-	2	2
1	-	-	-	-	3	2
2	2	-	-	-	4	3
2	1	-	-	-	5	3
1	2	-	-	-	6	3
1	1	-	-	-	7	3
2	2	2	-	-	8	4
2	2	1	-	-	9	4
2	1	2	-	-	10	4
2	1	1	-	-	11	4
1	1	1	-	-	12	4
1	1	2	-	-	13	4
1	2	1	-	-	14	4
1	2	2	-	-	15	4
2	2	2	2	-	16	5
2	2	2	1	-	17	5
2	2	1	2	-	18	5
2	2	1	1	-	19	5
2	1	2	2	-	20	5
2	1	2	1	-	21	5
2	1	1	2	-	22	5
2	1	1	1	-	23	5
1	1	1	2	-	24	5
1	1	1	1	-	25	5
1	1	2	2	-	26	5
1	1	2	1	-	27	5
1	2	1	2	-	28	5
1	2	1	1	-	29	5
1	2	2	2	-	30	5
1	2	2	1	-	31	5
-	-	-	-	-	32	1
4	-	-	-	-	33	2
3	-	-	-	-	34	2
4	4	-	-	-	35	3
4	3	-	-	-	36	3
3	4	-	-	-	37	3
3	3	-	-	-	38	3
4	4	4	-	-	39	4
4	4	3	-	-	40	4
4	3	4	-	-	41	4
4	3	3	-	-	42	4
3	3	3	-	-	43	4
3	3	4	-	-	44	4
3	4	3	-	-	45	4
3	4	4	-	-	46	4
4	4	4	4	-	47	5
4	4	4	3	-	48	5
4	4	3	4	-	49	5
4	4	3	3	-	50	5
4	3	4	4	-	51	5
4	3	4	3	-	52	5
4	3	3	4	-	53	5
4	3	3	3	-	54	5
3	3	3	4	-	55	5
3	3	3	3	-	56	5
3	3	4	4	-	57	5
3	3	4	3	-	58	5
3	4	3	4	-	59	5
3	4	3	3	-	60	5
3	4	4	4	-	61	5
3	4	4	3	-	62	5

Note: Allocations: 1 = 5/5; 2 = 9/1; 3 = 6/4; 4 = 8/2

Table A2: Scenarios of Recipients

Voter[1]	Voter[2]	Voter[3]	Voter[4]	Voter[5]	Scenario	Outcome
1	1	1	1	1	1	5/5
1	1	1	1	2	2	5/5
1	1	1	2	1	3	5/5
1	1	2	1	1	4	5/5
1	2	1	1	1	5	5/5
2	1	1	1	1	6	5/5
1	1	1	2	2	7	5/5
1	1	2	2	1	8	5/5
1	1	2	1	2	9	5/5
1	2	1	2	1	10	5/5
1	2	1	1	2	11	5/5
1	2	2	1	1	12	5/5
2	1	1	1	2	13	5/5
2	1	1	2	1	14	5/5
2	1	2	1	1	15	5/5
2	2	1	1	1	16	5/5
1	1	2	2	2	17	9/1
1	2	1	2	2	18	9/1
1	2	2	2	1	19	9/1
1	2	2	1	2	20	9/1
2	2	2	1	1	21	9/1
2	2	1	1	2	22	9/1
2	1	2	2	1	23	9/1
2	1	2	1	2	24	9/1
2	1	1	2	2	25	9/1
2	2	1	2	1	26	9/1
1	2	2	2	2	27	9/1
2	1	2	2	2	28	9/1
2	2	1	2	2	29	9/1
2	2	2	1	2	30	9/1
2	2	2	2	1	31	9/1
2	2	2	2	2	32	9/1
3	3	3	3	3	33	6/4
3	3	3	3	4	34	6/4
3	3	3	4	3	35	6/4
3	3	4	3	3	36	6/4
3	4	3	3	3	37	6/4
4	3	3	3	3	38	6/4
3	3	3	4	4	39	6/4
3	3	4	4	3	40	6/4
3	3	4	3	4	41	6/4
3	4	3	4	3	42	6/4
3	4	3	3	4	43	6/4
3	4	4	3	3	44	6/4
4	3	3	3	4	45	6/4
4	3	3	4	3	46	6/4
4	3	4	3	3	47	6/4
4	4	3	3	3	48	6/4
3	3	4	4	4	49	8/2
3	4	3	4	4	50	8/2
3	4	4	4	3	51	8/2
3	4	4	3	4	52	8/2
4	4	4	3	3	53	8/2
4	4	3	3	4	54	8/2
4	3	4	4	3	55	8/2
4	3	4	3	4	56	8/2
4	3	3	4	4	57	8/2
4	4	3	4	3	58	8/2
3	4	4	4	4	59	8/2
4	3	4	4	4	60	8/2
4	4	3	4	4	61	8/2
4	4	4	3	4	62	8/2
4	4	4	4	3	63	8/2
4	4	4	4	4	64	8/2

Note: Allocations: 1 = 5/5; 2 = 9/1; 3 = 6/4; 4 = 8/2

A.2 Behavioral Results

Table A3: Average sanction points for different voter roles and voter positions

		Fair Outcome - Voter Position						Unfair Outcome - Voter Position					
		1	2	3	4	5	Total	1	2	3	4	5	Total
Both	Minority 1	-0.65	-0.64	-0.71	-0.79	-1.08	-0.71	0.45	0.52	0.55	0.71	1.20	0.57
	Minority 2	.	-0.53	-0.53	-0.55	-0.50	-0.52	.	0.45	0.20	0.18	0.25	0.24
	Initiator	0.20	0.32	0.23	.	.	0.23	-0.98	-1.02	-1.33	.	.	-1.02
	Majority 2	.	0.18	0.27	0.22	.	0.22	.	-0.53	-0.69	-0.75	.	-0.72
	Pivotal	.	.	0.27	0.22	0.41	0.30	.	.	-0.65	-0.73	-1.02	-0.82
	Majority 4	.	.	.	0.17	0.12	0.14	.	.	.	-0.37	-0.29	-0.32
	Majority 5	.	.	.	.	0.18	0.18	.	.	.	.	-0.38	-0.38
Punish	Minority 1	-0.74	-0.74	-0.94	-1.13	-1.03	-0.85	-0.03	-0.04	-0.06	-0.03	-0.03	-0.04
	Minority 2	.	-0.62	-0.77	-0.63	-0.86	-0.75	.	-0.02	-0.09	0	-0.04	-0.04
	Initiator	-0.02	-0.03	-0.2	.	.	-0.03	-1.11	-1.21	-1.8	.	.	-1.18
	Majority 2	.	0	-0.06	-0.05	.	-0.04	.	-0.87	-1.37	-1.28	.	-0.99
	Pivotal	.	.	-0.06	-0.06	0	-0.04	.	.	-1.54	-1.4	-1.64	-1.52
	Majority 4	.	.	.	-0.05	-0.03	-0.04	.	.	.	-0.63	-0.7	-0.67
	Majority 5	.	.	.	.	-0.03	-0.03	.	.	.	.	-0.52	-0.52
Reward	Minority 1	0.15	0.2	0.25	0.29	0.41	0.22	0.69	0.88	0.95	0.7	1.31	0.84
	Minority 2	.	0.12	0.12	0.12	0.18	0.14	.	0.62	0.68	0.72	0.5	0.62
	Initiator	0.88	1.04	1.12	.	.	0.94	0.09	0.19	0.17	.	.	0.12
	Majority 2	.	1.09	1.02	1.13	.	0.97	.	0.12	0.1	0.12	.	0.12
	Pivotal	.	.	1.01	1.14	1.36	1.19	.	.	0.09	0.12	0.16	0.13
	Majority 4	.	.	.	0.54	0.54	0.54	.	.	.	0.12	0.12	0.12
	Majority 5	.	.	.	.	0.53	0.53	.	.	.	.	0.14	0.14

A.3 Individual Regressions - Theoretical Framework

Table A4: Punishment Treatment

Punishment Points	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Choice Unfair	1.013***								0.770***		0.797***
	(0.123)								(0.177)		(0.176)
Outcome Unfair		0.529***							0.0336		-0.0530
		(0.111)							(0.0993)		(0.0724)
Intention Unkind			0.947***						0.0509		-0.0613
			(0.124)						(0.130)		(0.138)
Initiator Unfair				0.705**					0.296		0.190
				(0.244)					(0.233)		(0.231)
Pivotal Unfair					1.090***				0.643**		0.598**
					(0.240)				(0.211)		(0.208)
BF Responsibility Unfair						2.698***				2.716***	0.0306
						(0.431)				(0.486)	(0.349)
Ex-ante Engl Resp Unfair							2.362***			-0.0473	-0.0416
							(0.382)			(0.341)	(0.221)
Ex-post Engl Resp Unfair								1.175***		0.0151	0.346*
								(0.194)		(0.220)	(0.144)
Constant	0.0377	0.279***	0.153**	0.474***	0.435***	0.274***	0.271***	0.285***	0.0272	0.275***	0.0442
	(0.0615)	(0.0557)	(0.0511)	(0.0244)	(0.0240)	(0.0431)	(0.0442)	(0.0428)	(0.0591)	(0.0441)	(0.0562)
Observations	9,600	9,600	9,600	9,600	9,600	9,600	9,600	9,600	9,600	9,600	9,600
R-squared	0.181	0.050	0.154	0.032	0.076	0.136	0.118	0.111	0.207	0.136	0.210
Number of Subjects	30	30	30	30	30	30	30	30	30	30	30

Note: Fixed effects regression with punishment points as dependent variable (ranges from 0 to 7). *BF Responsibility Unfair* represents a voters' share in the probability increase of an unfair outcome. *Ex-ante and Ex-post Engl Responsibility Unfair* represents a voters' causal responsibility for an unfair event. Robust standard errors in parentheses clustered at the subject level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A5: Reward Treatment

Reward Points	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Choice Fair	0.758***								0.451**		0.468***
	(0.134)								(0.123)		(0.119)
Outcome Fair		0.408***							0.0797		0.0231
		(0.0787)							(0.0606)		(0.0410)
Intention Kind			0.750***						0.261**		0.187
			(0.135)						(0.0942)		(0.0964)
Initiator Fair				0.459***					0.0250		-0.0282
				(0.0898)					(0.0644)		(0.0639)
Pivotal Fair					0.744***				0.281*		0.265**
					(0.153)				(0.115)		(0.0825)
BF Responsibility Fair						1.833***				1.241***	0.0736
						(0.348)				(0.303)	(0.258)
Ex-ante Engl Resp Fair							1.830***			-0.277	-0.583*
							(0.345)			(0.251)	(0.256)
Ex-post Engl Resp Fair								0.890***		0.494***	0.439**
								(0.163)		(0.131)	(0.132)
Constant	0.144*	0.318***	0.213***	0.476***	0.448***	0.339***	0.317***	0.312***	0.119	0.312***	0.130
	(0.0672)	(0.0393)	(0.0556)	(0.00898)	(0.0153)	(0.0348)	(0.0387)	(0.0385)	(0.0653)	(0.0390)	(0.0638)
Observations	9,280	9,280	9,280	9,280	9,280	9,280	9,280	9,280	9,280	9,280	9,280
R-squared	0.143	0.042	0.136	0.019	0.050	0.101	0.087	0.095	0.161	0.108	0.165
Number of Subjects	29	29	29	29	29	29	29	29	29	29	29

Note: Fixed effects regression with reward points as dependent variable (ranges from 0 to 7). *BF Responsibility Fair* represents a voters' share in the probability increase of a fair outcome. *Ex-ante and Ex-post Engl Responsibility Fair* represent a voters causal responsibility for a fair event. Robust standard errors in parentheses clustered at the subject level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A6: Both Treatment - Punishment Points

Punishment Points	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Choice Unfair	0.729***								0.481***		0.514***
	(0.120)								(0.103)		(0.104)
Outcome Unfair		0.353***							0.00766		-0.0751*
		(0.0744)							(0.0491)		(0.0356)
Intention Unkind			0.709***						0.177*		0.0941
			(0.124)						(0.0768)		(0.0734)
Initiator Unfair				0.714**					0.344		0.252
				(0.246)					(0.240)		(0.247)
Pivotal Unfair					0.498**				0.149		0.202*
					(0.142)				(0.117)		(0.0929)
BF Responsibility Unfair						1.733***				1.066*	-0.344
						(0.322)				(0.435)	(0.344)
Ex-ante Engl Resp Unfair							1.774***			-0.654	-0.181
							(0.325)			(0.340)	(0.315)
Ex-post Engl Resp Unfair								0.852***		0.683*	0.451***
								(0.161)		(0.258)	(0.118)
Constant	0.0269	0.215***	0.0988	0.320***	0.342***	0.218***	0.196***	0.185***	0.0245	0.192***	0.0330
	(0.0600)	(0.0372)	(0.0510)	(0.0246)	(0.0142)	(0.0322)	(0.0358)	(0.0390)	(0.0582)	(0.0373)	(0.0578)
Observations	9,600	9,600	9,600	9,600	9,600	9,600	9,600	9,600	9,600	9,600	9,600
R-squared	0.133	0.031	0.122	0.046	0.022	0.082	0.077	0.088	0.148	0.093	0.151
Number of Subjects	30	30	30	30	30	30	30	30	30	30	30

Note: Fixed effects regression with punishment points as dependent variable (ranges from 0 to 7). *BF Responsibility Unfair* represents a voters' share in the probability increase of an unfair outcome. *Ex-ante and Ex-post Engl Responsibility Unfair* represents a voters' causal responsibility for an unfair event. Robust standard errors in parentheses clustered at the subject level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A7: Both Treatment - Reward Points

Reward Points	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Choice Fair	0.309***								0.374***		0.372***
	(0.0613)								(0.0898)		(0.0879)
Outcome Fair		0.0152							-0.112*		-0.0995**
		(0.0506)							(0.0443)		(0.0323)
Intention Kind			0.247***						-0.0220		-0.00178
			(0.0551)						(0.0432)		(0.0381)
Initiator Fair				0.0902					-0.0395		-0.0439
				(0.0510)					(0.0339)		(0.0328)
Pivotal Fair					0.150				0.0147		0.0504
					(0.0823)				(0.0681)		(0.0543)
BF Responsibility Fair						0.379				0.569**	-0.129
						(0.188)				(0.173)	(0.152)
Ex-ante Engl Resp Fair							0.322			-0.449*	-0.283
							(0.181)			(0.179)	(0.195)
Ex-post Engl Resp Fair								0.172		0.136	0.154
								(0.0919)		(0.0740)	(0.114)
Constant	0.0196	0.166***	0.0722**	0.165***	0.159***	0.136***	0.138***	0.138***	0.0545	0.140***	0.0507
	(0.0307)	(0.0253)	(0.0227)	(0.00510)	(0.00823)	(0.0188)	(0.0205)	(0.0192)	(0.0300)	(0.0200)	(0.0282)
Observations	9,600	9,600	9,600	9,600	9,600	9,600	9,600	9,600	9,600	9,600	9,600
R-squared	0.062	0.000	0.038	0.002	0.005	0.010	0.007	0.008	0.070	0.011	0.071
Number of Subjects	30	30	30	30	30	30	30	30	30	30	30

Note: Fixed effects regression with reward points as dependent variable (ranges from 0 to 7). *BF Responsibility Fair* represents a voters' share in the probability increase of a fair outcome. *Ex-ante and Ex-Post Engl Responsibility* represents a voters causal responsibility for a fair event. Robust standard errors in parentheses clustered at the subject level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.4 Robustness Check Cluster Analysis

Results: Classification - Responsibility Measures.

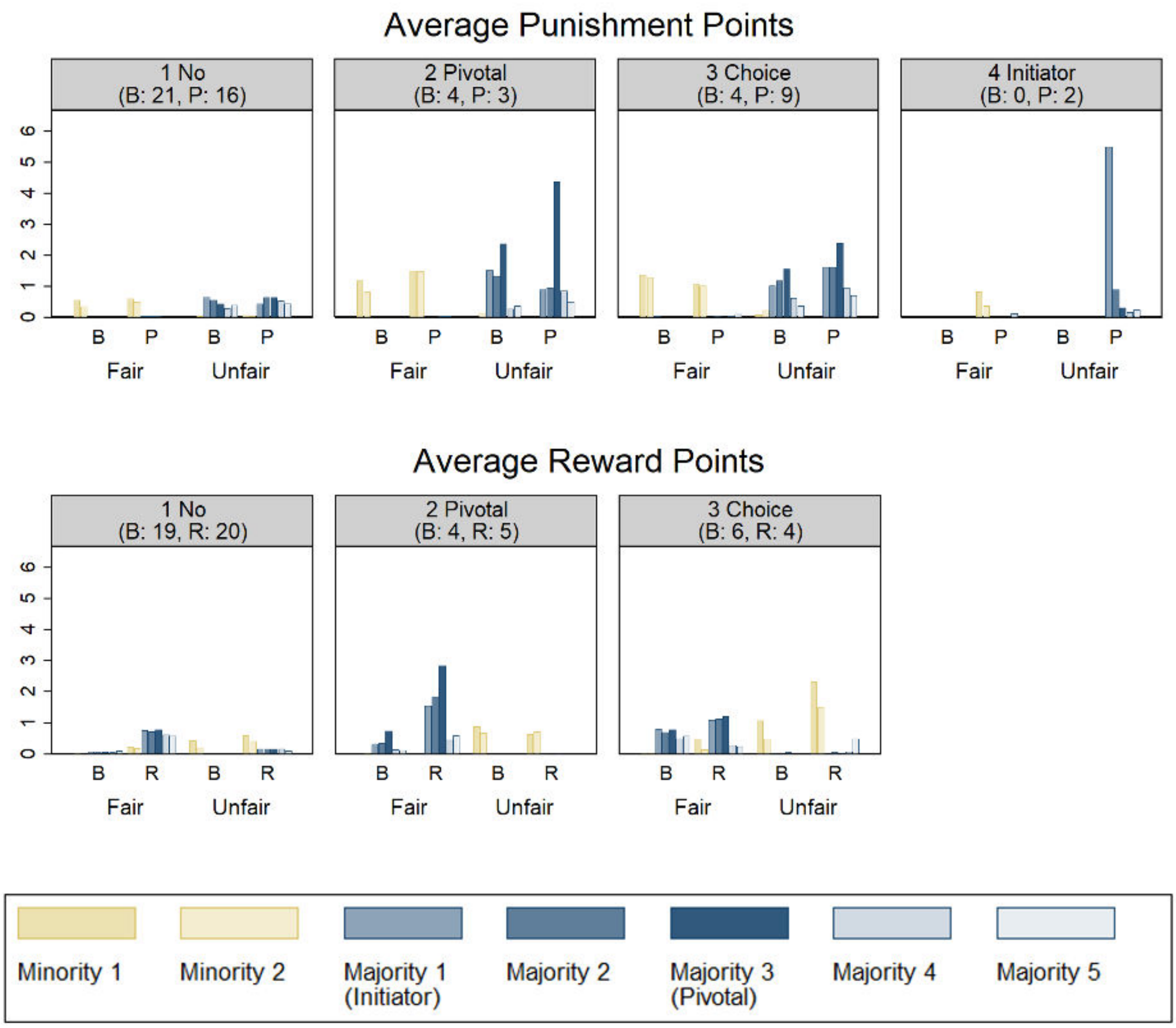


Figure A3: Cluster analysis: Punishment and reward patterns across treatments based on responsibility measures

Note: The figure shows the average punishment and reward points used in each cluster for fair and unfair outcomes across treatments. Hereby, the punishing patterns in the treatments *Punishment* and *Both* are presented in the upper part of the figure, while the reward patterns in the treatments *Reward* and *Both* are presented in the lower part of the figure. B = *Both*, P = *Punishment*, R = *Reward*, n = number of recipients in the corresponding cluster.

In order to perform a robustness check of the findings presented in Section 4.1.2 and to see whether the responsibility measures *BF Responsibility* and *Engl Responsibility* capture the same dynamics of how responsibility is attributed, the same cluster analysis is performed based on these measures. The procedure stays the same as in the first cluster analysis despite of the set of variables that is used for the individual regressions. Here, we only use the measures *BF Responsibility* and *Engl Ex-Ante Responsibility* and *Engl Ex-Post Responsibility* separated by fair and unfair.[18]

In line with the first cluster analysis, we find that subjects in Cluster 1 do barely punish or reward, subjects in Cluster 2 show a preference for punishing / rewarding the pivotal voter and subjects in Cluster 3 mainly punish unfair / reward fair choices. The fourth punishment cluster captures the same pattern as in the fourth punishment cluster in the first cluster analysis.

The two cluster analyses, where the first one is based on the individual sanctioning motives while the second is only based on the responsibility measures, manage to identify very similar behavioral patterns among the subjects. Table A8 compares the classification of subjects into clusters based on the two variable sets.

Table A8: Classification of subjects based on two cluster analyses with different variable sets (individual motives and responsibility measures) in treatment *Punishment* and *Reward*

		Cluster analysis: responsibility measures								
		Punishment					Reward			
		No	Piv	Choice	Initiator	Total	No	Piv	Choice	Total
Cluster analysis: individual measures	No	12	0	3	0	15	14	0	0	14
	Piv	0	2	3	0	5	0	5	0	5
	Choices	4	1	3	0	8	5	0	4	9
	Initiator	0	0	0	2	2	-	-	-	-
	Total	16	3	9	2	30	19	5	4	28

The classification of subjects in the *Reward* treatment based on the two variable sets is very similar. Although there are 5 subjects who are classified as using little or no reward when using responsibility measure and otherwise are classified as focusing on the choices when using individual motives, the remaining subjects are classified into the exact same clusters. Despite some deviations, the majority of subjects in the *Punishment* treatment are classified into the same clusters when comparing the two cluster analyses.

The results of the second cluster analysis show that the responsibility measures by Bartling et al. (2015) and Engl (2018) identify the same patterns as a heterogeneity analysis based on individual motives. We argue that despite serving as a robustness check, the responsibility measure capture the same dynamics as the individual measures.

[18]Note that we select the same number of clusters as in the first cluster analysis to assess the robustness of our results. However, this is not always the optimal number of clusters. The result of the second cluster analysis with the optimal number of clusters can be found in Appendix A.4.1.

A.4.1 Cluster analysis based on responsibility measures with optimal number of clusters

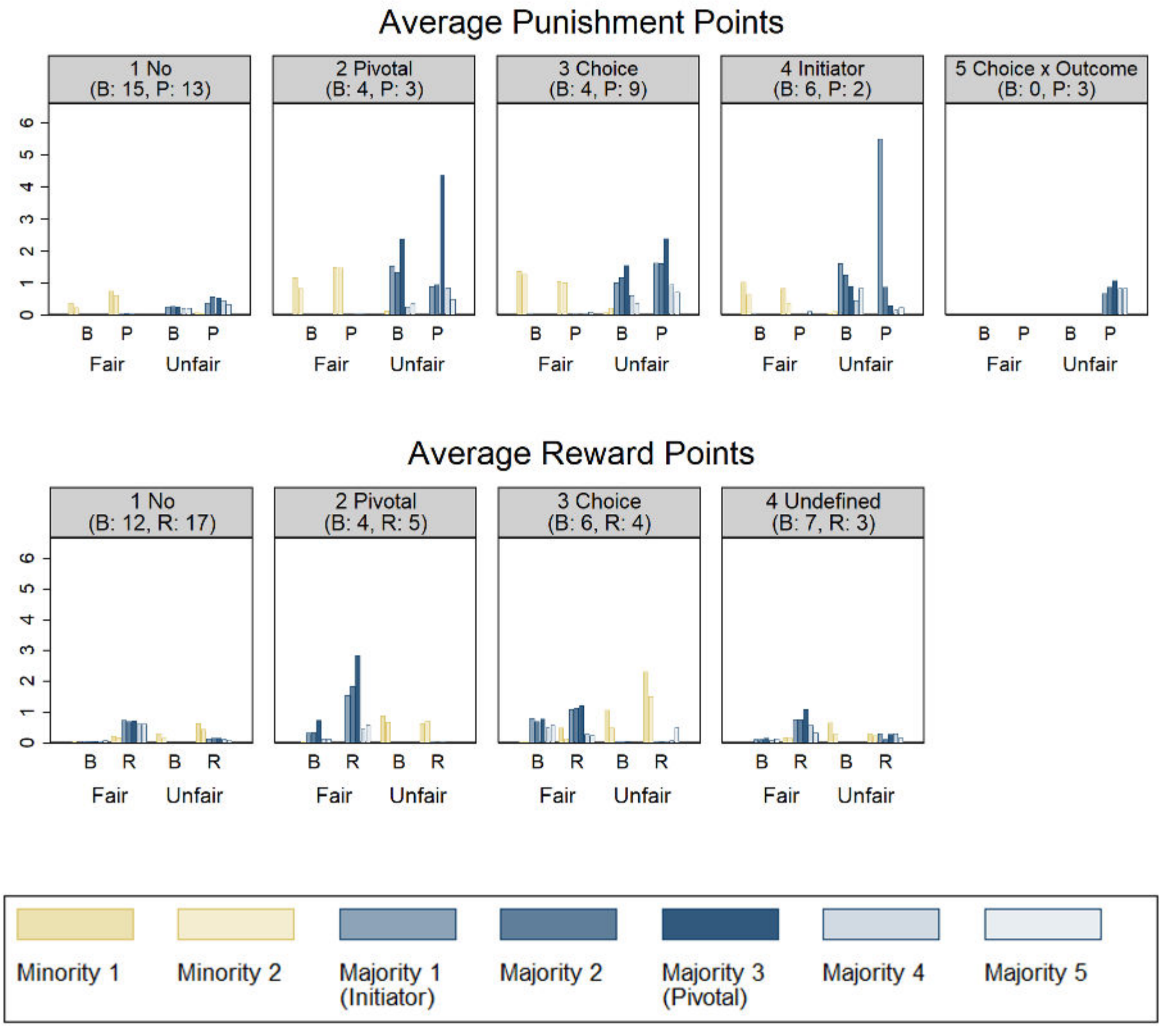


Figure A4: Cluster analysis: Punishment and reward patterns across treatments based on responsibility measures with optimal number of clusters.

Note: The figure shows the average punishment and reward points used in each cluster for fair and unfair outcomes across treatments. Hereby, the punishing patterns in the treatments *Punishment* and *Both* are presented in the upper part of the figure, while the reward patterns in the treatments *Reward* and *Both* are presented in the lower part of the figure. B = *Both*, P = *Punishment*, R = *Reward*, n = number of recipients in the corresponding cluster.

A.5 Eye–Tracking Regressions - Second Decision Screen

Table A9: Share of Fixations - Importance of Position and Saliency on the Second Screen

	Punishment	Reward	Both
Voter Position 2	0.042	0.096***	0.029
	(0.022)	(0.014)	(0.025)
Voter Position 3	0.122**	0.123***	0.047
	(0.035)	(0.012)	(0.036)
Voter Position 4	0.032	0.050**	-0.024
	(0.031)	(0.017)	(0.035)
Voter Position 5	-0.050	-0.007	-0.075*
	(0.028)	(0.015)	(0.030)
Only One Salient	0.050	0.099**	0.180***
	(0.024)	(0.032)	(0.030)
Neighbor of Only One Salient	-0.014	0.015	-0.002
	(0.012)	(0.015)	(0.015)
Minority	-0.015	0.018	0.017
	(0.019)	(0.016)	(0.019)
Constant	0.174***	0.134***	0.188***
	(0.024)	(0.012)	(0.024)
Observations	4270	4030	4040
R-squared	0.054	0.055	0.068
Number of Subjects	28	27	24

Note: Fixed effects regression with share of fixations as dependent variable fixing on Voter Position 1. *Voter Position 2-5* indicate the position of the voter in the decision scenario. *Only One Salient* is a dummy variable indicating if the voter is the only salient voter. *Neighbor of Only One Salient* is a dummy variable that takes the value 1 if the voter is next to the only salient voter. *Minority* is a dummy variable indicating if the voter is part of the minority. Robust standard errors in parentheses clustered at the subject level.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A10: Share of Fixations - Impact of Sanction Motives on the Second Screen

	Punishment	Both		Reward	Both
Choice Unfair	0.0512	0.0149	Choice Fair	0.0859***	-0.0511*
	(0.0291)	(0.0188)		(0.0230)	(0.0192)
Intention Unkind	0.0433*	0.0458*	Intention Kind	0.0426*	-0.0202
	(0.0176)	(0.0201)		(0.0193)	(0.0152)
Initiator Unfair	0.0330	0.0372*	Initiator Fair	0.0450*	0.0163
	(0.0299)	(0.0179)		(0.0199)	(0.0227)
Pivotal Unfair	-0.00325	-0.0129	Pivotal Fair	-0.00763	0.0350
	(0.0196)	(0.0183)		(0.0225)	(0.0180)
Controls	Yes	Yes	Controls	Yes	Yes
Constant	0.0932**	0.131***	Constant	0.0288	0.209***
	(0.0317)	(0.0233)		(0.0202)	(0.0269)
Observations	4,270	4,040	Observations	4,030	4,040
R-squared	0.087	0.085	R-squared	0.119	0.082
Number of Subjects	28	24	Number of Subjects	27	24

Note: OLS fixed effects regression with share of fixations of the second screen as the dependent variable fixing on Voter Position 1. The controls include variables for *voter position 2-5, saliency and minority.* Robust standard errors in parentheses clustered at the subject level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.6 Instructions for Participants

The following pages display the original instructions (translated into English) used in the experiment for the recipients. The corresponding instructions for the voters are almost the same as for the recipients.

General explanations for participants B

We warmly welcome you to this economic study.

If you read the following explanations carefully, then - depending on your decisions and the decisions of the other participants - you can earn money in addition to the **5 euros** that you receive as a show-up fee. It is therefore very important that you read these explanations carefully. If you have any questions, please address them to us.

During the study, you are not allowed to talk to the other participants in the study. Failure to comply with this rule will result in exclusion from the study and all payments.

The study consists of 2 parts. The instructions for both parts can be found on the following pages. Part 1 consists of 62 decisions and part 2 of 2 decisions. At the end, a draw is made to determine whether a decision from Part 1 or Part 2 is payout relevant. The probability that a decision of part 1 is relevant for payment is 96% and the probability that a decision from part 2 is relevant for payment is 4%.

During the study, we do not speak of euros, but of points. So, your total income is first calculated in points. The total number of points you earn during the study will then be converted into euros at the end, where the following conversion rate applies

1 point = 3 euros.

At the end of today's study, we will pay you the number of points earned during the study plus €5 in **cash** for showing up. The people who use an eye tracker will receive another €5 extra as compensation for the inconvenience. You will not find out which people use an eye tracker.

On the following pages we explain the exact procedure of the study.

The study

At the beginning, you will be randomly and anonymously assigned to nine other people who are also participating in this study. Neither before nor after the study will you learn the identity of the nine people assigned to you. The nine people assigned to you will also not learn anything about your identity.

In this study, there are two types of participants: Participant A and Participant B.

You are a participant B.

Each group consists of five participants A and five participants B. Thus, five participants A and four additional participants B are assigned to you.

The study consists of 64 scenarios spread over two parts. This means that each participant makes 64decisions. At the end of the experiment, one scenario is randomly drawn to be realized for the payout.

Part 1

In Part 1, the five Participants A decide by majority vote how 50 points will be divided between the five Participants A and the five Participants B.

Here, there are two different situations in which Participant A can choose from two different allocations.

<u>**Situation 1:**</u>

Participants A must choose between the following two possible allocations:

- **Allocation 1:** Participants A get 9 points each and participants B get 1 point each.
- **Allocation 2:** Participants A and Participants B get 5 points each.

The following table gives you an overview of the two allocations that Participants A must decide between.

	A	A	A	A	A	B	B	B	B	B
Allocation 1	9	9	9	9	9	1	1	1	1	1
Allocation 2	5	5	5	5	5	5	5	5	5	5

Situation 2:

Participants A must choose between the following two possible allocations:

- **Allocation 1:** Participants A get 8 points each and participants B get 2 points each.
- **Allocation 2:** Participants A get 6 points each and participants B get 4 points each.

The following table gives you an overview of the two allocations that Participants A must decide between.

	A	A	A	A	A	B	B	B	B	B
Division 1	8	8	8	8	8	2	2	2	2	2
Division 2	6	6	6	6	6	4	4	4	4	4

Regardless of the situation, the distribution that receives the majority of votes by participants A will be implemented. So, if three or more of the participants A decide for Allocation 1, then Allocation 1 is implemented. If three or more of the participants A decide for Allocation 2, then Allocation 2 is implemented.

Abstentions are not possible. Each participant A must vote either for Allocation 1 or for Allocation 2.

The voting procedure:

Participants A vote on the allocations **one after the other.**

1. The participant A who decides first is participant **A1**.
2. Participant A, who decides second, is participant **A2**. Participant A2 observes how participant A1 has decided before making his own decision.
3. Participant A, who is the third to decide, is participant **A3.** Participant A3 observes how participants A1 and A2 have decided before making his own decision.
4. Participant A, who decides fourth, is participant **A4. Participant** A4 observes how participants A1, A2 and A3 have decided before making his own decision.
5. The participant A who decides last is participant **A5**. Participant A5 observes how participants A1, A2, A3 and A4 have decided before making his own decision.

The allocation that at least three of the five Participants A opt for will be implemented.

The voting result is therefore fixed as soon as three participants A have decided on the same allocation.

The decisions of the participants B:

Participants B learn not only the result of the vote, but also how each individual participant A decided. Participants B therefore learn how first participant A1, then participant A2, A3, A4 and finally participant A5 decided.

After the voting outcome of participants A is determined, a participant B is drawn, who receives an extra point.

This participant B then has the option to give or deduct points from each participant A by giving up the extra point. The selected participant B can give and/or deduct up to 7 points in total from the participants A.

Any (whole) number of points between 0 and 7 points can be given or taken away. As soon as at least 1 point is given or taken away from participants A, the extra point is deducted from participant B. **So, the costs to give or deduct points is always equal to the extra point, regardless of the number of given or deducted points.**

For example, if the drawn participant B wants to deduct 7 points from participant A3, the payout of participant A3 will be reduced by **7 points**. The payoff of participant B is equal to the payoff after the voting by participant A before getting the extra point.

For example, if the drawn participant B wants to give 5 points to participant A1 and deducts 1 point from participant A4, participant A1's payout will increase by **5 points** and participant A4's payout will decrease by **1 point,** but participant B will not receive any extra point because he used it to deduct or give points.

The only restrictions on giving or deducting points are that no more than a total of 7 points can ever be deducted or given, and that no more points can ever be deducted from a participant A than he or she received in the distribution chosen by the majority.

If in situation 1 the distribution 2 (5/5) is realized, because four participants A decided in favor and one participant A decided against, the participant B, who is drawn and uses his extra point, can deduct a **maximum of 5 points from** a participant A, because they did not get more by the voting outcome.

As soon as participant B has decided to give/take points from participant A, the extra point is used. It does not matter whether or not points are actually given or taken away in the end, since the extra point is used directly with the decision to give/take points.

Participants A will only find out at the end whether and how many points Participant B has given them/subtracted.

Part 2

After all Participants A have played through the possible scenarios of majority decisions, two more decisions follow for Participants A. In these two rounds, both situations (Situation 1: Allocations (9/1; 5/5) or Situation 2: Allocations (8/2; 6/4)) **are** played out again, **with no majority decision and no reward or punishment.** This means that although all participants A decide to allocate in both situations, only one participant A is randomly selected, whose decision is realized for all participants A and participants B in his group. Participants B do not get an extra point afterwards and therefore cannot give/take any points.

The course of the experiment

In this study, a total of 64 scenarios are possible for you as participant B, depending on which situation (situation 1: Allocations (9/1; 5/5) or situation 2: Allocations (8/2; 6/4)) you are in and which allocations participants A1-A5 have chosen. The 64 scenarios correspond to all possible decision combinations of participants A, regardless of whether the participants are in part 1 or part 2.

Before you find out if you are by random decision the participant B who can give or subtract points from the participants A, we ask you to **indicate for each possible scenario how you would decide.**

You have the option of taking a break after 17 scenarios and removing your head from the chin rest. If you want to continue, you must position the head on the chin rest again and click on "Next". Please do not take longer than 1 minute for this break.

At the end of the experiment, it is first randomly determined whether a scenario of the majority decisions from part 1 or the decisions from part 2 is played out. Then, the situation (situation 1: Allocations (9/1; 5/5) or situation 2: Allocations (8/2; 6/4)) is determined by a dice roll. Then, Participants A are randomly assigned their role (A1-A5) and a Participant B is randomly selected to give/take points. If you get the role of this participant B, the corresponding scenario, which you already faced, is payoff-relevant and thus your choice for this scenario. Thus, only one scenario will be paid out in the end.

Below you will find a numerical example to illustrate one scenario of this study. Afterwards you will find the control questions, which also serve for your understanding. The information on these questions is not relevant for your payout.

Please complete the comprehension questions and raise your hand afterwards so we can review them.

A numerical example

Part 1 and Situation 1 were determined randomly.

Participant A1 chooses Allocation 1 (9/1).

Participant A2 observes A1's decision and then chooses Allocation 2 (5/5).

Participant A3 observes the decisions of A1 and A2 and then chooses Allocation 1 (9/1).

Participant A4 observes A1's, A2's, and A3's decisions and then chooses Allocation 1 (9/1).

The outcome of the vote has now already been determined. The result is Allocation 1 (9/1).

Participant 5 observes the decisions of his predecessors and then chooses Allocation 2 (5/5).

All participants B are informed about the chosen allocation of each participant A.

Participant B1 is drawn by lot and can give and subtract points from participant A.

B1 gives **3 points to** participant A2 and draws **4 points** from participant A4.

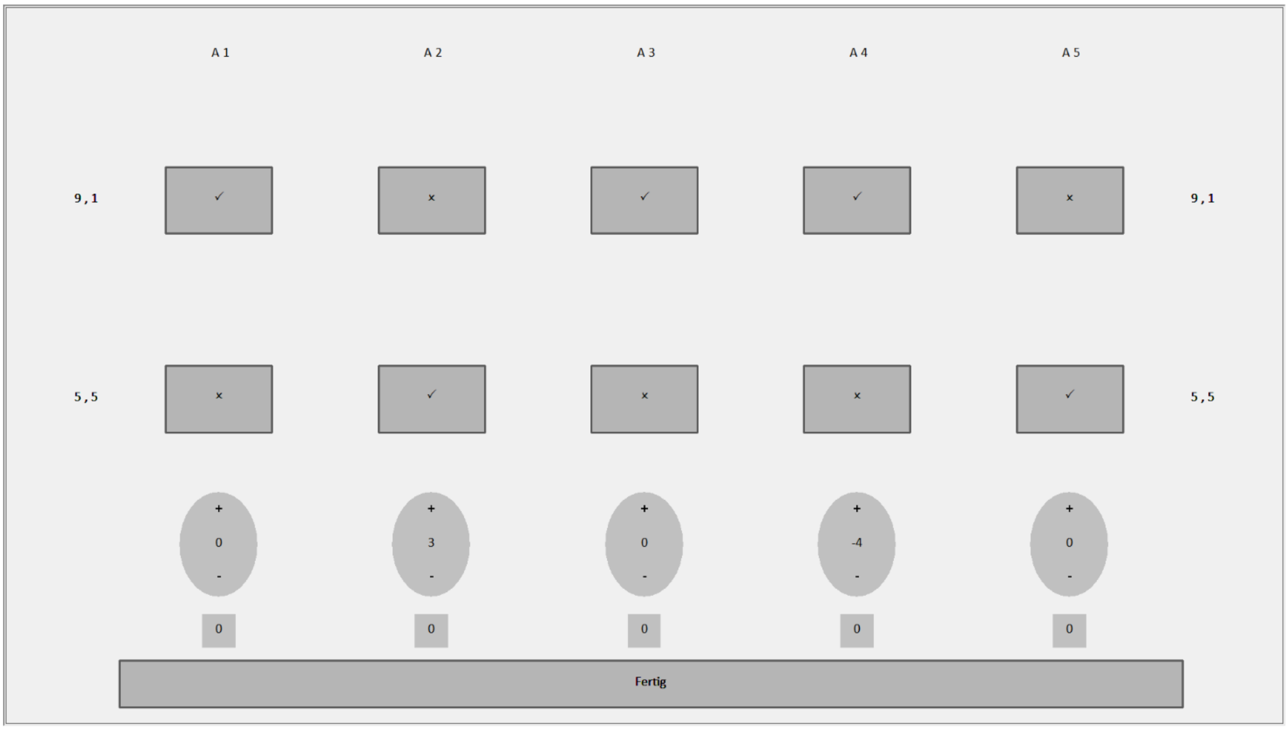
This results in the following payments:

	Decisions of the participants A	Points from the distribution	Points by participant B	Extra point?	Final points
A1	Allocation 1	9	0	-	9
A2	Allocation 2	9	3	-	12
A3	Allocation 1	9	0	-	9
A4	Allocation 1	9	-4	-	5
A5	Allocation 2	9	0	-	9
B1	-	1	-	used	1
B2	-	1	-	-	1
B3	-	1	-	-	1
B4	-	1	-	-	1
B5	-	1	-	-	1

Your process on the computer

The numerical example just shown on page 7 will look like this on your screen. Note that you have to make a decision for each scenario, but in the end only one participant B will be selected from your group whose decision will be implemented.

First of all, for each scenario you can see how all participants A have decided. Now you can decide whether to use your extra point to give or deduct points from participants A, or leave the scenario without action.



If you decided to use your extra point, you can give/take away up to 7 points to the participants A.

Control questions

Question 1:

A scenario from **Part 1 - Situation 1** (9/1; 5/5) has been determined. Fill in the gray fields.

	Decisions of the participants A	Resulting points?	Bonus points through B	Extra point?	Payouts?
A1	Allocation 1		-2	-	
A2	Allocation 1		0	-	
A3	Allocation 2		2	-	
A4	Allocation 2		2	-	
A5	Allocation 1		-1	-	
B1	-		-	used	
B2	-		-	-	
B3	-		-	-	
B4	-		-	-	
B5	-		-	-	

Question 2:

Related to question 1: What would be the maximum number of points that participant B1 could deduct participant A5?

Response:

Question 3:

A scenario from **part 2 - situation 1** (9/1; 5/5) was determined. At the end, a participant A was drawn, who chose Allocation 1.

How many points will each participant A receive? Answer:

How many points will each participant B receive? Answer:

Can a participant B give/take points in part 2? Answer:

Question 4:

Does participant B have to use his extra point in every scenario from part 1?

Response:

Question 5:

What is the probability that a scenario from Part 2 will be determined for payout?

Response:

THURGAU INSTITUTE
OF ECONOMICS
at the University of Konstanz

Hafenstrasse 6
CH-8280 Kreuzlingen

T +41 (0)71 677 05 10
F +41 (0)71 677 05 11

info@twi-kreuzlingen.ch
www.twi-kreuzlingen.ch