

A Social Scoring Systems for Behavioral Regulation: An Experiment on the Role of Transparency in Determining Perceptions and Behaviors - Appendix

This document contains details about the experimental sessions, as well as the scoring mechanism. We report the instructions and intervention messages in the treatments, as well as the survey scales used to retrieve people’s perceptions. Illustrations of people’s perceptions and their final standings in the respective treatments are displayed.

Table 2: Overview of sessions (mean values, standard deviations in brackets). In each treatment, four experimental sessions were run. T: Transparent, O: Opaque, B: Baseline

	T	O	B
Participants	13.75 (9.11)	15 (7.26)	12.5 (8.35)
Payment in Euro	11.65 (1.84)	16.80 (5.93)	14.37 (3.34)
Rounds	16.75 (2.75)	22 (9.35)	21.5 (3.42)
Duration in minutes	37.0 (8.8)	46.7 (15.3)	31.0 (7.2)

Level of trustworthiness	Monetary range	Penalty per MU retained
Very untrustworthy	$b < \frac{r}{3}$	4 points
Untrustworthy	$\frac{r}{3} \leq b < \frac{r}{2}$	2 points
Trustworthy	$\frac{r}{2} \leq b < e$	1 point
Very trustworthy	$b \geq e$	no penalty

Table 3: Point penalties applied to the score based on different levels of (un)trustworthiness. The penalties are cumulative for each monetary unit (MU) retained relative to the payoff-equalizing amount. For example, receiving 20 MUs and sending back 0 MUs leads to a rounded point penalty of 38 points: 4 points for each MU retained up to one-third of the received amount, 2 points for each MU retained between one-third and one-half, and 1 point for each MU retained between one-half and the payoff-equalizing amount. Here, b is the amount sent back, r is the amount received, and e is the payoff-equalizing amount.

Standing	Score range
A	1000-985
B	984-970
C	969-955
D	954-940
E	939-0

Table 4: Mapping of scores to standings.

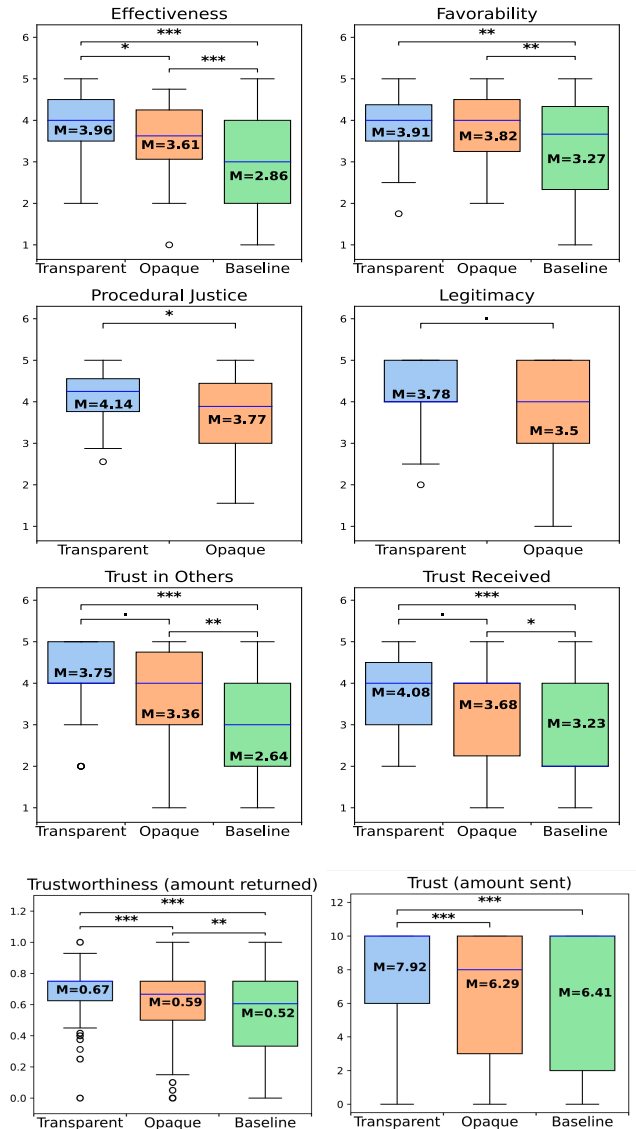


Figure 3: Boxplots of behaviors and perceptions across treatments. The brackets indicate significant differences between treatments (one-sided Mann-Whitney U tests); $p < .001$ (***), $p < .01$ (**), $p < .05$ (*), $p < .1$ (\cdot). Mean values (M) are displayed.

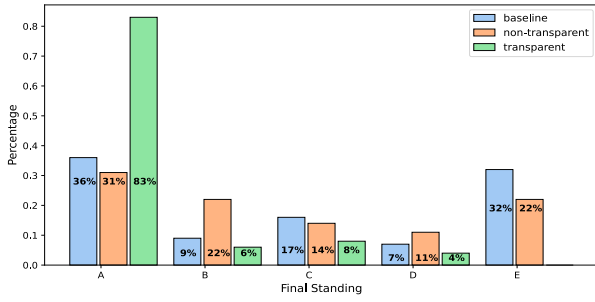


Figure 4: Overview of the composition of final standings as a percentage of the participants. For comparison, the hypothetical standings in the baseline treatment are added.

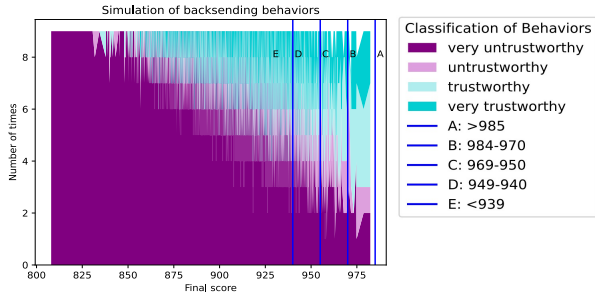


Figure 5: Simulation of final scores after 10 rounds as a second mover using the point-penalty mechanism in Table 3. The amount received was drawn from the distribution of the first-movers' sending behaviors, the return ratio from the distribution of second-movers' return behaviors in (Berg, Dickhaut, and McCabe 1995). The x-axis indicates the final score, the y-axis indicates the number of times a behavior of a certain trustworthiness level occurred. The plot displays the 90th percent quantile of the simulations. The highest-scoring simulation resulted in a score of 982, but still showed untrustworthy behaviors. We, therefore, set the cutoff for the highest standing at 985. The lowest-scoring simulated participant in the 90th percent quantile showed a score of 935. To ensure that some sort of downgrading occurred, we set the cutoff for lowest standing at 940, and set the remaining ranges evenly.

Table 5: Survey Scales

Perceived Procedural Justice

- I feel the institution acted for the well-being of the community.
 - I think the institution made decisions that were good for everyone in the community.
 - I feel the institution did what is necessary to help the community.
 - I think the institution treated community members with dignity and respect.
 - I understood the rules and methods the institution used to decide on my standing.
 - I feel that I knew how my behavior would impact my standing.
 - I think the mechanism the institution used to assess my standing was fair.
 - I was able to influence the data that the institution considered to evaluate my standing.
 - I feel able to influence my standing such that I am satisfied with it.
-

Perceived Legitimacy

- I felt a moral duty to do what the institution asked participants of the community to do.
 - I generally support how the institution acted.
 - The institution had the same sense of right and wrong as I do.
 - I could trust the institution to make the right decisions.
-

Perceived Favorability

- Generally speaking, I am satisfied with how the experiment went for me.
 - I generally liked how I was treated by other community members.
 - Having the institution in the community was favorable to me.
 - I am satisfied with my final standing.
-

Perceived Effectiveness

- I think the institution successfully influenced people in my community to behave in fair manner.
 - I think the institutions made good decisions regarding the standing of the people in my community.
 - Having the institution in my community was useful to me.
-

Perceived Trust

- I had the impression that the people with whom I interacted trusted me.
 - I could trust the people with whom I interacted.
-

Table 6: Instructions in the respective treatments (Tr.). T: Transparent, O: Opaque, B: Baseline

Tr.	General Instructions
B	<p>Intervention: To achieve fairness, the institution will provide you with a summary of your past interactions from round number 2 on. In addition, in case the distribution of the payoffs between the first and the second mover is highly unequal, the institution will remind you to treat your match partners in a fair manner.</p>
O	<p>Evaluation: To achieve fairness, the institution will evaluate your decisions as a second mover. The institution assigns a standing to each participant. The standing ranges from A to E, where A is the highest standing. The standing is updated according to the distribution of payoffs between the first mover and the second mover. Everyone starts with the highest standing. In case the distribution of payoffs between the first and the second mover is highly unequal, the institution might update the standing of the second mover. Once the standing is updated, (e.g. from A to B), the previous standing cannot be reached again.</p> <p>Visibility of Evaluation: In case you are the first mover, you can request your match partner’s standing before you decide how much of your endowment to send. When you are in the role of a second mover, your match partner can equally request your standing before deciding how much to send to you.</p>
T	<p>Evaluation: To achieve fairness, the institution will evaluate your decisions as a second mover. The institution assigns a score to each participant, ranging from 1000 to 0. 1000 is the best score, 0 is the worst score. Every participant in the community starts with a score of 1000. The score is updated according to the distribution of payoffs between the first mover and the second mover. In case the distribution of payoffs between the first and the second mover is highly unequal, points may be subtracted. Based on your score, the institution assigns you to a standing. This standing ranges from A to E, and is done according to the following table [Table 4 was displayed]. Once your standing is updated, (e.g. from A to B) the previous standing cannot be reached again.</p> <p>Visibility of Evaluation: In case you are the first mover, you can request your match partner’s standing before you decide how much of your endowment to send. You can only request your match partner’s standing, not the score. When you are in the role of a second mover, your match partner can equally request your standing before deciding how much to send to you. Your match partner can request only your standing, your score. Your score is only visible to you.</p>

Table 7: Feedback messages in the respective treatments (Tr.). B: Baseline, O: Opaque, T: Transparent

Tr.	Feedback Message
B	<p>Feedback: The institution found that you did not treat your match partner as you have been encouraged to. The institution decided to remind you to respect the instructions. You can consult them below.</p>
O	<p>Feedback: The institution found that you did not treat your match partner as you have been encouraged to. Therefore, the institution decided to update your standing. Your current standing is [...].</p>
T	<p>Feedback: The institution found that you did not treat your match partner as you have been encouraged to. The institution decided to lower your score to [...]. Therefore, the institution also updated your standing. Your current standing is [...].</p>