Surprising Gifts

Theory and Laboratory Evidence

Kiryl Khalmetski,[a] Axel Ockenfels,[b,*] Peter Werner[c]

4 May 2015

Accepted for publication in *Journal of Economic Theory*

Abstract

People do not only feel guilt from not living up to others' expectations (Battigalli and Dufwenberg (2007)), but may also like to exceed them. We propose a model that generalizes the guilt aversion model to capture the possibility of positive surprises when making gifts. A model extension allows decision makers to care about others' attribution of intentions behind surprises. We test the model in a series of dictator game experiments. We find a strong causal effect of recipients' expectations on dictators' transfers. Moreover, in line with our model, the correlation between transfers and expectations can be both positive and negative, obscuring the effect in the aggregate. Finally, we provide evidence that dictators care about what recipients know about the intentions behind surprises.

Keywords: guilt aversion, surprise seeking, dictator game, consensus effect

JEL-Codes: C91, D64

[*] Corresponding author.

[a] University of Frankfurt, Senckenberganlage 31, D-60325 Frankfurt am Main, Germany (e-mail: kiryl.khalmetski at uni-koeln.de). *Present address:* University of Cologne, Albertus-Magnus-Platz, D-50923 Köln, Germany, tel.: +492214701476.

[b] University of Cologne, Albertus-Magnus-Platz, D-50923 Köln, Germany (e-mail: ockenfels at uni-koeln.de), tel.: +492214705761.

[c] University of Cologne, Albertus-Magnus-Platz, D-50923 Köln, Germany (e-mail: peter.werner at uni-koeln.de), tel.: +492214704354.

1. INTRODUCTION

Models of guilt aversion assume that people feel guilt from not living up to others' expectations (Battigalli and Dufwenberg (2007), henceforth "BD"). Yet, it appears plausible that some people do not only suffer from negative surprises, but may also get pleasure from positive surprises (e.g., Mellers et al. (1997)). We thus propose a generalized model of guilt aversion by incorporating the notion that people may care for both positive and negative surprises when making gifts.[1] In case of dictator games, our model implies that the dictator may experience a utility loss from falling short of the recipient's expected transfer, and a utility gain from exceeding it, both being a potential motivation to transfer money to the recipient. Moreover, our model predicts a positive correlation between transfers and expectations for dictators who want to avoid negative surprises, yet a *negative* correlation for dictators who have a relatively strong preference for creating positive surprises. The underlying rationale for the negative correlation is that there is more room to positively surprise a recipient with lower expectations; that is, the marginal utility gain from a positive surprise is increased by lowered expectation.

We test the model's predictions in a series of dictator game experiments and find strong support. Moreover, we show that our data reconcile seemingly conflicting evidence from previous studies on guilt aversion.

Our Experiment 1 is designed to investigate the prediction that dictator transfers can both decrease and increase with the recipient's expectation, depending on the weight put on positive and negative surprises, respectively. We find a strong causal effect of recipients' expectations on individual dictator transfers. The effect is obscured on the aggregate level because, as suggested by our model, dictators differ in how they react to the recipients' expectations.

Our evidence sheds light on the controversy about whether others' expectations directly affect social behavior or not. By eliciting subjects' beliefs about the expectations of interaction partners (second-order beliefs, SOBs), several studies

[1] Our research is part of the literature that is devoted to people's concern about beliefs *per se*, independently of the material outcome (Geanakoplos et al. (1989), Bénabou and Tirole (2006), Andreoni and Bernheim (2009)). The framework of dynamic psychological games (Battigalli and Dufwenberg (2009)) incorporates many of these earlier approaches, including the notion that people suffer from guilt when they disappoint what they think are other players' expectations.

detected a positive relation between beliefs and observed behavior. The first study along these lines was conducted by Dufwenberg and Gneezy (2000). In an experimental “lost-wallet” game, a player could either take an amount of money or pass the decision to a second player who then had to decide on how to split a larger amount between the two. The authors find that the decisions of the second player were positively correlated with their beliefs about what the first players expected from them as a transfer. In a study by Charness and Dufwenberg (2006), subjects who held significantly higher beliefs about their transaction partner’s expectation were also more trustworthy. This is in line with several other experiments that have found positive correlations between subjects’ self-reported beliefs and observed decisions.[2]

However, more recently, others have argued that correlations between self-reported SOBs and choices may be confounded by the false consensus effect (Ross et al. (1977)): the SOB might be biased towards one’s own choice. If this is the case, observed correlations between actions and beliefs are no conclusive evidence for beliefs causally affecting behavior.

To address concerns about consensus effects, Ellingsen et al. (2010, henceforth "EJTT") *induced* SOBs in an experimental dictator game by disclosing the first-order beliefs (FOBs) elicited from recipients to dictators. While it can be criticized that recipients did not know that their FOBs would be transmitted to dictators, and that this was known by the latter (see our discussion below), the design allowed to establish a direct *causal* influence of SOBs on giving. Yet, no correlation was found between induced SOBs and actual behavior. The authors thus concluded that the empirical support for guilt aversion might be limited and partly confounded by the false consensus effect.[3]

---

[2] For experimental evidence on the impact of belief-dependent preferences in trust, dilemma and principal-agent games see also Guerra and Zizzo (2004), Falk and Kosfeld (2006), Reuben et al. (2009), Dufwenberg et al. (2011) and Charness and Dufwenberg (2011). Vanberg (2008) investigated potential reasons behind the positive effect of promises on trustworthy behavior found in Charness and Dufwenberg (2006) and concluded that preferences for promise-keeping rather than preferences for meeting expectations might be the predominant driver of the results. With respect to dictator and ultimatum games, the willingness of some subjects to exploit information asymmetries between themselves and recipients suggests that behavior depends on beliefs (see, for example, Mitzkewitz and Nagel (1993), Güth et al. (1996), Güth and Huck (1997), Dana et al. (2007), Andreoni and Bernheim (2009), Grossman (2010, 2014), Ockenfels and Werner (2012), Taubinsky (2012), and Cappelen et al. (2013)).

[3] See the references in EJTT and, for more recent, mixed laboratory evidence on guilt aversion, Bellemare et al. (2011, 2014), Attanasi et al. (2014), Kawagoe and Narita (2014).

Our Experiment 1 closely follows the EJTT experiment design. Indeed, we replicate all main results of EJTT's dictator treatment, including the lack of correlation between transfers and induced SOBs *on the between-subject level*. Moreover, we provide further evidence for the confounding role of the false consensus effect. However, in addition to EJTT, we elicit transfers for many expectation levels of the recipient. This allows us to investigate the different individual patterns of behavior that we expect to see based on our model. The *within-subject* data show that many subjects do systematically condition transfers on the recipients' expectations, suggesting that both guilt aversion and a preference for creating positive surprises are relevant. Yet, because in line with our model we observe both positive and negative within-subject correlations of transfers with expectations, no such correlation can be identified for the aggregated data.

In Experiment 2, we take a complementary approach to study the performance of our model in the laboratory. Here, unlike in Experiment 1, we are interested in creating a situation in which the comparative static prediction of our model is unambiguous in the sense that it does not depend on the weight put on negative and positive surprises in the dictator's utility function. At the same time, we are interested to learn more about the nature of the dictator's motivation for surprising. More specifically, BD introduced two models of guilt aversion. One is "simple guilt" and refers to a player who cares about the extent to which he lets another player down. The second model, "guilt from blame", assumes that a player cares about others' inferences regarding the extent to which he is willing to let them down (i.e., inferences about his intentions). We formulate our model to capture the potential role of 'intentional surprise'. The model predicts that if the recipient's inference about the dictator's intention is ambiguous, the latter has weaker incentives both to avoid guilt and to positively surprise the recipient, and should in turn transfer less. Importantly, this effect is predicted for both relatively guilt-averse and surprise-seeking dictators.[4] To test this prediction we introduce an experiment design, which manipulates the

[4] With respect to the terminology concerning the dictators' preferences, we note that both negative and positive deviations from expected transfers are surprising to the recipient, of course. Yet, dictators who are guilt-averse aim at *reducing* the element of negative surprise by living up to the recipient's expectations. On the contrary, those dictators who positively deviate from the expectation aim at increasing surprise and are thus termed as "surprise seekers".

recipient's inference about the dictator's intentions by making the recipient either aware or unaware about the fact that the dictator's SOB was induced. The results of the experiment support our model's prediction, and show that dictators care about the recipients' attribution of their intentions behind surprises.

EJTT's experiment design and findings constitute a simple, well-known and influential challenge for guilt aversion in general, and for our model in particular. This is why we test our hypotheses in the same environment, utilizing the same method of controlling SOBs, as used in EJTT. However, we caution that some aspects of EJTT's experiment design are controversial, because the way information is withheld from subjects might lead to a loss of control similar to deceiving subjects (although subjects are not literally deceived but surprised). In EJTT's experiments, recipients are not aware about the fact that their beliefs will be transmitted so that dictators can condition transfers on them.[5] Also, dictators might get suspicious when learning, before making their choices, that recipients were not informed about all strategically relevant aspects of the decision situation. This might create the impression that there are also possibly other aspects of the design that are withheld from the dictators. Overall, the procedures might lead to a general suspicion among participants that seemingly simple decisions may have unforeseen consequences, which eventually distorts the decisions made in the experiment. Thus, it is important to not only closely relate our findings to EJTT's results (as we do in Experiments 1 and 2), but to also conduct robustness checks which mitigate some of the potential problems inherent in EJTT's design.

Specifically, in our Experiments 3 and 4, all subjects have the same information about information flows at every stage of the experiment. At the same time, recipients explicitly decide whether or not to disclose their beliefs, while, using a novel design (described in Section 4), we make sure that these beliefs have not been strategically distorted. All central results from the first two studies can be replicated with our new experiment design.

Section 2 presents our generalized model of surprising, describes the experimental design to investigate both guilt aversion and surprise seeking, analyzes the data and compares them to related results in the literature. Section 3 extends the model of

[5] Recipients were debriefed about actual information flows *after* they submitted their beliefs.

surprising to capture the effect of the recipient's inference about the dictator's intention, and presents the design and the results of the second experiment. Section 4 describes our robustness checks and Section 5 concludes.

2. A MODEL OF SURPRISING OTHERS AND EXPERIMENT 1

2.1. Model

The dictator game is a useful starting point for our application of psychological game theory to demonstrate the impact of surprises (and intentions behind surprises) on social behavior, because it abstracts away from potentially confounding strategic or reciprocal interaction. Assume that dictator *i* divides an amount normalized to 1 between himself and recipient *j*, who holds an *ex ante* expectation E_j about her payoff from the dictator's transfer t_i. Applying the benchmark model of guilt aversion of BD, the utility of the dictator from transferring t_i is

$$u_i(t_i, E_j) = m_i(1-t_i) - \eta_i \max\left\{0, E_j - m_j(t_i)\right\}, \tag{1}$$

where $m_i(\cdot)$ is the standard utility of money, further assumed to have conventional properties $m_i(0)=0$, $m_i'(\cdot)>0$ and $m_i''(\cdot)<0$,[6] $\max\left\{0, E_j - m_j(t_i)\right\}$ is the dictator's level of guilt from falling below the recipient's expectation, and η_i is a coefficient reflecting guilt sensitivity.

In the formalization by BD, guilt is strictly positive only for transfers strictly below expectations. That is, only negative surprises matter. However, based on the idea that people like pleasant surprises, it appears reasonable that both negative and positive deviations from the recipient's expectation directly enter the dictator's utility function. More specifically, we assume that dictators do not only suffer from

[6] The non-linearity of m_i is needed to guarantee the existence of interior solutions for optimal transfers (see Lemma 2 in Appendix A).

negatively surprising the recipient, but also derive utility from positively surprising her.[7]

Moreover, while some applied models of guilt aversion (e.g., Charness and Dufwenberg (2011) and Beck et al. (2013)) take the point expectation as the recipient's reference point, other models of reference-dependent preferences exploit the whole distribution of beliefs (a reference lottery; e.g., Köszegi and Rabin (2006)). That is, the *ex post* outcome is compared with all outcomes in the support of the reference lottery weighted by the corresponding *ex ante* probabilities.[8] Following this approach, we assume that the reference point of the recipient, against which the surprise is evaluated, is given by a probability distribution of possible outcomes (i.e., the reference point is stochastic). As we show in Appendix A, the distribution-wise representation of beliefs allows that the marginal surprise (and hence, the optimal transfer) continuously changes with expectations, without precluding discrete jumps.[9]

We further denote the cumulative distribution function (cdf) of the FOB of the recipient as H_j, with the corresponding probability density function (pdf) h_j. The above assumptions lead to the following extension of the dictator's utility function (1):

$$u_i(t_i \mid h_j) = m_i(1 - t_i) + S_i(t_i \mid h_j), \tag{2}$$

with

[7] A related psychological game approach is Ruffle (1999) who formulates a model in which positive surprising is only possible in mixed-strategy equilibria. His approach differs from ours in that our model mainly focuses on the analysis of optimal pure strategies, allowing for the possibility of non-equilibrium beliefs. Geanakoplos et al. (1989), too, present an example of a psychological game where the recipient likes being positively surprised.

[8] The *ex post* outcome is perceived as a gain relative to lower possible outcomes, and as a loss relative to higher possible ones. Larsen et al. (2004), among others, provide psychological evidence that people experience different feelings of both pleasure and pain while simultaneously comparing the realized outcome with, respectively, lower and higher counterfactuals.

[9] An alternative approach to get the optimal transfer continuously change with beliefs would be to introduce some nonlinearity directly into the guilt function, which, however, would require additional assumptions on its functional form. Another possible way would be to keep a point-wise recipient's expectation as her reference point, but to introduce some uncertainty for the dictator regarding it, which would imply that the SOB of the dictator is stochastic, as stated in our subsequent model. However, such approach would result in a deterministic dictator's reference point if a moment of the recipient's FOB distribution is exogenously signaled to the dictator, which is the focus of our laboratory analysis.

$$S_i(t_i \mid h_j) = \alpha_i \int_0^{t_i} (t_i - x) h_j(x) dx - \beta_i \int_{t_i}^{1} (x - t_i) h_j(x) dx \, . \tag{3}$$

Hereafter, S_i is referred to as the surprise function. The first term in the surprise function represents the dictator's utility from positive surprises (when $x < t_i$), while the second one represents the disutility from negative surprises (when $x > t_i$). The stochastic reference point is the distribution of FOB, given by the pdf h_j. Correspondingly, the scalar $\alpha_i \geq 0$ denotes the propensity to make positive surprises (surprise seeking), and the scalar $\beta_i \geq 0$ corresponds to the propensity to avoid negative surprises (guilt aversion). These propensities are not necessarily equal. In order to simplify the exposition, we assume that the value of surprise for a particular belief x and transfer t_i (the term weighted by $h_j(x)$) is linear in $x - t_i$.

In what follows, we make the following assumptions on the utility function:

A1. $\alpha_i \neq 0 \vee \beta_i \neq 0$.

A2. $m_i'(1 - t_i) \geq \dfrac{\alpha_i + \beta_i}{2}$ for any $t_i \in [0,1]$.

Assumption A1 rules out the trivial case when the dictator does not have any belief-dependent preferences.[10] Assumption A2 states that the marginal monetary cost of giving is larger than the average sensitivity to positive and negative surprises. For example, if $\alpha_i = 0$ and $\beta_i > 0$, Assumption A2 requires that for a decrease in the negative surprise term by 1 Euro a dictator is willing to pay at most 2 Euro.[11] The

[10] When both coefficients are 0, the dictator chooses zero transfers for any beliefs. In principle, one could also add other motives, not related to beliefs, such as inequality aversion to the utility function (Fehr and Schmidt (1999) and Bolton and Ockenfels (2000)). Then, there can be positive equilibrium transfers even if $\alpha_i = \beta_i = 0$. Yet, the transfers will be independent of beliefs (which is the null hypothesis in the subsequent Experiment 1). Assumption A1 is also needed to ensure the strict comparative statics prediction of Proposition 2.

[11] In this case the decrease in the negative surprise term by 1 Euro yields an increase in the total utility equal to $\beta_i \Delta$, and the payment of additional 2 Euro results in a loss of at least $m_i'(w_0) 2\Delta$, where Δ corresponds to 1 Euro in terms of normalized amounts and w_0 is the initial level of wealth. Assumption A2 implies that the loss is weakly higher than the gain. Technically, the assumption is needed to prove Proposition 2.

assumption is well in line with existing estimations of the quantitative effect of guilt aversion (Bellemare et al. (2011)).

Regarding the information structure of the game, we assume that the FOB of the recipient is unknown to the dictator. Yet, he observes an informative signal θ_j about the FOB, which is equal to the median of the FOB distribution. Then, his SOB is characterized by the conditional cdf $H_{ij}(x \mid \theta_j) = E_i[H_j(x) \mid \theta_j]$ with the corresponding conditional pdf $h_{ij}(\cdot \mid \theta_j)$.[12] We emphasize that we do not require that the recipient's individual beliefs correspond to a rational expectations equilibrium (as it is assumed in BD). Rather, we treat θ_j as exogenous to the model. Mutual consistency of beliefs and behavior is rejected by numerous dictator game experiments, including EJTT (see their Figure 1, which reveals significant heterogeneity in beliefs about average dictator transfers). Of course, our modeling does neither exclude the possibility that beliefs are consistent with behavior, nor that average beliefs are consistent with average behavior (as roughly observed by Selten and Ockenfels (1998), among others).

Further, we do not explicitly model how the dictator forms his SOB as the expectation of the recipient's FOB conditional on the obtained signal. Instead, we implement a reduced-form model, assuming only that a higher signal leads to a higher SOB in the sense of first-order stochastic dominance (FOSD).

A3. The SOB conditional on a higher signal (strictly) first-order stochastically dominates the SOB conditional on a lower signal:

$H_{ij}(x \mid \theta''_j) < H_{ij}(x \mid \theta'_j)$ if and only if $\theta''_j > \theta'_j$ for any $x \in (0,1)$ and $\theta'_j, \theta''_j \in [0,1]$.[13]

Finally, we impose some smoothness on the cdf function H_{ij}:

---

[12] The alternative assumption that θ_j is the *mean* of the FOB distribution would require a slight change of Assumption A2, keeping the underlying intuition behind this assumption unchanged. The original model of BD does not require a signaling parameter θ_j, because here beliefs are assumed to be mutually consistent in equilibrium.

[13] $H_{ij}(x \mid \theta_j)$ does not depend on θ_j at $x = 0$ and $x = 1$ (being always equal to 0 and 1, respectively).

A4. $H_{ij}(x|\theta_j)$ is continuously differentiable on $[0,1]\times[0,1]$.

We denote further the values of the surprise and the expected utility functions given transfer t_i and signal θ_j as $S_i(t_i,\theta_j)$ and $U_i(t_i,\theta_j)$, respectively.

Let us now consider the optimal strategy of the dictator. For simplicity, we assume that the dictator plays a pure strategy conditional on the signal θ_j, i.e., the dictator's chosen transfer can be represented as a function t_i^* of the signal, so that

$$t_i^*(\theta_j) \in \arg\max_{t_i} U_i(t_i,\theta_j). \tag{4}$$

Proposition 1 shows how the optimal transfer $t_i^*(\theta_j)$ depends on the dictator's signal θ_j about the recipient's FOB.[14]

Proposition 1. *For relatively guilt-averse dictators (with* $\alpha_i < \beta_i$*) the optimal transfer is increasing in the signal* θ_j*, while it is decreasing in* θ_j *for relatively surprise-seeking dictators (with* $\alpha_i > \beta_i$*). The increase (decrease) is strict if* $0 < t_i^*(\theta_j) < 1$.

Proof. See Appendix A. ■

Intuitively, this result stems from the fact that the reference point of the recipient is stochastic. Indeed, if the recipient's expectations increase (i.e., θ_j gets larger), and the whole distribution of beliefs is shifted to the right, then it follows that a smaller part of the beliefs distribution is exceeded by a given transfer. Consequently, the marginal gain from positively surprising the recipient gets smaller since a smaller mass of beliefs is affected by the surprise (see (3)). This eventually leads to a lower

[14] The proof of Proposition 1 does not require Assumption A2 and the fact that θ_j is the median of the recipient's FOB distribution, but is consistent with them. In case $\alpha_i = \beta_i$ the surprise term converges to the representation with point-wise beliefs: $S_i = -\alpha_i(E_{j|\theta_j} - t_i)$. Then, the optimal transfer does not depend on θ_j. The consistency of the interior solution condition $0 < t_i^*(\theta_j) < 1$ with both $\alpha_i > \beta_i$ and $\alpha_i < \beta_i$ is established by Lemma 2 in Appendix A.

transfer if the concern for positive surprises is relatively more important for the dictator $(\alpha_i > \beta_i)$. The opposite holds if the concern for negative surprises dominates.

2.2. Experiment design and hypotheses

Our Experiment 1 is designed to test whether heterogeneity in belief-dependent preferences (guilt aversion and surprise seeking) contributes to explaining dictator game behavior. As explained before, Experiment 1 closely follows the design of the dictator game experiment by EJTT, who informed dictators about recipients' expectations before the transfer had to be chosen. The only exception is that we used the strategy method (Selten (1967)) in our experiment in order to elicit dictator transfers *conditional* on the recipient's potential expectations, whereas EJTT employed the direct-response method. We choose this design because, as we show below, this allows us not only to replicate the EJTT findings (as a robustness check), but also to detect a potential heterogeneity in dictators' transfers as a function of expectations, as suggested by our model, and so to provide a new interpretation of EJTT's dictator game data.[15]

In our experiment, each dictator had to divide 14 Euro between himself and a randomly matched anonymous recipient. Before observing the actual amount sent to her, the recipient was asked to provide a guess for the average transfer in the population. In order to stay close to EJTT's design, the guess closest to the average transfer was rewarded with an additional bonus of 8 Euro.[16] Before the guess of the recipient was revealed to the corresponding dictator, he was asked to indicate his transfer conditional on all possible guesses rounded to 50 cents. Guesses higher than 9 Euro were grouped into a single category.[17] This way, we collected 19 choices per

---

[15] We use the term "strategy method" in a broad sense; dictators do not condition their decisions on other subjects' actions – as it is typically the case with the strategy method – but on beliefs about expected transfers (see concluding section for a discussion of this method in our context). Attanasi et al. (2014) and Bellemare et al. (2014) applied similar methods in the context of dictator and trust games to detect behavioral patterns in line with guilt aversion.

[16] As in EJTT, recipients did not know (at the time of submission of the guess) that their beliefs will be transmitted to dictators, and dictators were aware of this fact.

[17] This was done to not bother subjects with reflecting on unlikely guesses. In fact, guesses higher than 9 Euro were chosen by less than 5% of the recipients, and, likewise, transfers higher than 9 Euro were realized in less than 5% of cases.

experimental subject. After that, the conditional transfer, which corresponded to the actual expectation level of the matched recipient, was implemented and paid out.

The experiment was conducted as a classroom experiment among students of economics and business at the University of Frankfurt with a total of 386 students. The classroom was divided in two separate halves by a central aisle. Students sitting in the first half received the dictator's instructions, and students sitting in the second one received the recipient's instructions. Instructions can be found in Appendix E.

The recipient's guess is assumed to be close to the median of her FOB distribution,[18] and hence we also assume that it corresponds to signal θ_j in terms of our model. Then, the strategy method allows us to infer the mapping of signals θ_j to transfers $t_i^*(\theta_j)$ for each dictator. In other words, the design allows us to investigate a dictator's willingness to give as a function of the recipient's expectation. Our null hypothesis is that dictators do not have belief-dependent preferences, and hence transfers are independent of the recipients' guesses. However, if the dictator cares about positive or negative surprises of the recipient, Proposition 1 suggests that we should observe transfers which are (positively or negatively) correlated with expectations (our alternative hypothesis).

2.3. Experiment results

In total we obtained 3,629 observations for conditional transfers from 191 dictators (19 for each dictator), and 195 observations for recipients' guesses.[19] Our results are fully comparable to the results in the dictator game treatment of EJTT. The average actually realized transfer was 3.25 Euro. This was 23% of the endowment, which is approximately the same value as in EJTT, where it was 24% of the endowment. 28% of the dictators were not willing to transfer a strictly positive

---

[18] In the actual distribution of transfers which we observed in our experiment, the mean (3.25 Euro) was close to the median (3 Euro).

[19] Because of a matching error, we had four recipients more than needed; these recipients were paid according to the decisions made by a randomly chosen dictator of another pair. A single dictator provided only one conditional transfer, while leaving the fields for other conditional transfers blank. These blank fields were interpreted as zeros, though, none of our results are affected if we drop this dictator. Also, we always include transfers conditional on the highest guess category (larger than 9 Euro) in the data analyses, although our conclusions do not change if we exclude these values.

amount to their matched recipients. The corresponding value was 35% in EJTT. The average guess of recipients was 4.70 Euro. This was 34% of the endowment, compared to 32% in EJTT. Finally, EJTT emphasized that they did not find a correlation between guesses and transfers (Pearson correlation coefficient of −0.075, $p = 0.497$). In our experiment, the correlation between actually realized transfers and guesses, too, was not significantly different from zero (Pearson correlation coefficient of −0.017, $p = 0.821$).[20]

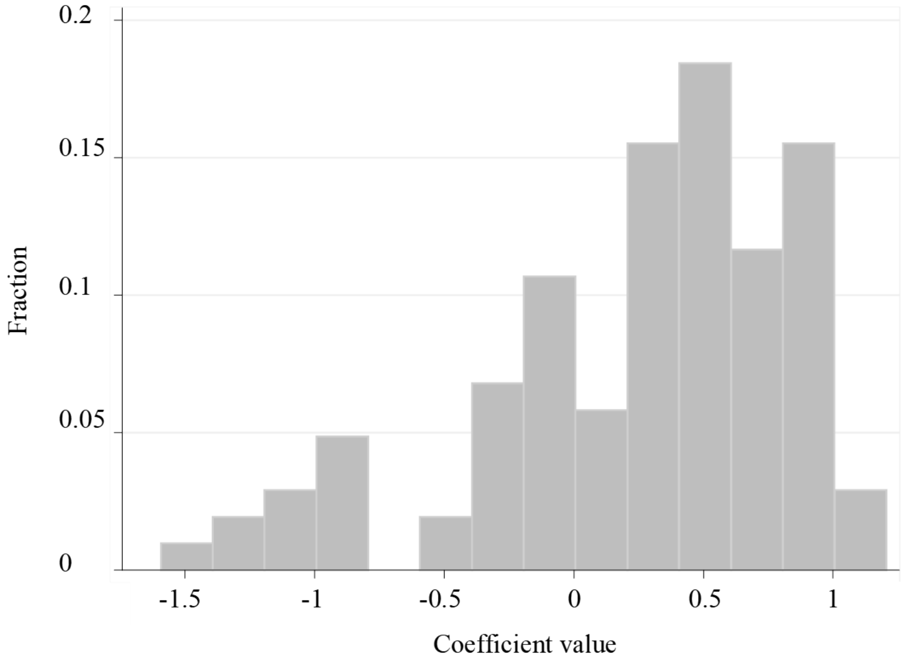
All these observations are in line with our null hypothesis. However, the within-subject data tell a more subtle story: 77.5% of the dictators changed their transfers in response to guesses at least once, and 53.9% of the dictators exhibited a within-subject correlation of transfers with guesses which is significant at the 5% level.[21] To check whether the observed patterns can be organized by a random process, we ran a Monte-Carlo simulation with 10,000 replications of random samples of transfers obtained by bootstrapping the original sample. On average, the share of significant within-subject correlations between transfers and guesses in random samples is just 3.7% with a standard deviation of 1.3%. None of the replications produced a sample with a share of significant correlations of more than 9.4%. We conclude that our observed share of 53.9% is the result of a systematic choice. This rejects all purely outcome-based models as an explanation of positive transfers, and demonstrates that many dictators care about recipients' beliefs.[22]

[20] Random matching is itself a stochastic process, and hence a single random matching may be not representative. Our within-subject design allows a more robust measure of the average correlation at the between-subject level by estimating correlation coefficients under different possible matchings between dictators and recipients. We performed a Monte-Carlo simulation of 10,000 random matching combinations between subjects, estimating correlation between transfers and guesses in each replication. The average Pearson correlation coefficient is with 0.102 ($p = 0.162$) a bit higher than the low coefficient that corresponds to the random matching used to pay out our subjects, but it is still not significantly different from zero.

[21] While the dictators with a significant correlation showed a strong tendency to either increase or decrease transfers in response to guesses, 30.1% of them changed behavior at least once in both directions. If we exclude these dictators from the sample, none of our results are qualitatively affected, except for the difference in strength between positive and negative individual correlations which becomes larger (see footnote 23).

[22] These and all subsequent results remain robust to excluding guesses higher than 7 Euro.

Fig. 1. The distribution of coefficients, significant at 5% level, estimated in within-subject regressions of transfers on guesses.



Consistent with Proposition 1, we also find that dictators differed qualitatively in how they responded to changes in recipients' expectations. Figure 1 shows the distribution of the statistically significant coefficients from regressing transfers on guesses for each dictator. According to Proposition 1 positive regression coefficients correspond to relatively guilt-averse dictators, while negative coefficients to relatively surprise-seeking dictators. Figure 1 shows that 69.9% of the coefficients are distributed to the right of zero. The asymmetry is statistically significant: a two-sided sign test strongly rejects the hypothesis that the median is equal to 0 ($p < 0.001$). Moreover, the average size of positive coefficients (0.58) is somewhat larger than the size of negative coefficients (0.53), although the difference is not statistically significant.[23] Overall, guilt aversion appears to be more prevalent than surprise seeking in our dictator game context. This seems consistent with reference-dependent preferences models (like Kahneman and Tversky (1979), Fehr and Schmidt (1999),

[23] $p = 0.123$, two-sided MWU-test. The difference in size between the average negative and positive coefficients becomes statistically significant, though, if we exclude those dictators who changed their transfers with expectations in both directions. Then, the average size of (significant) positive coefficients is 0.60, while it is 0.46 for negative coefficients ($p = 0.020$, two-sided MWU test). The relative share of dictators with significantly positive coefficients then remains almost unchanged (69.4%).

Köszegi and Rabin (2006)) along with empirical evidence (like Tversky and Kahneman (1992), Ockenfels et al. (forthcoming)), which suggest that falling below the respective reference standard generally has a larger effect on utility than a same-sized gain above the reference point.

At the same time, we find that the positive surprise side cannot be neglected either. For one, similar to EJTT (see their Figure 1), we find that 27.2% of all transfers submitted by dictators were strictly above guesses. Second, and more importantly, 30.1% of the dictators for whom we found a significant within-subject correlation between transfers and guesses exhibited a *negative* correlation. This corresponds to 16.2% of the total population.[24] This observation is inconsistent with pure guilt aversion, yet consistent with Proposition 1 of our generalized model.

2.4. False consensus

EJTT interpret the fact that there is no correlation between transfers and induced SOBs suggesting "*that consensus effects are responsible for a substantial fraction of the correlation between second-order beliefs and behavior in other studies*" (p. 101). In this view, the correlation between *self-reported* SOBs and behavior is not due to the SOBs causally affecting behavior, but rather due to a tendency of subjects to believe that others' behavior is similar to one's own behavior. However, our model suggests and our data show that EJTT's non-correlation result is caused by opposing causal effects of SOBs across individual dictators – guilt aversion and positive surprise seeking.[25]

At the same time, our data confirm EJTT's conjecture and others' findings that the false consensus is real. We measured the *ex ante* SOBs of dictators in a non-

[24] The highest share of dictators with a significantly negative within-subject correlation observed in 10,000 bootstrapped replications of random transfers was 6.3% (with 95% of replications yielding a share below 3.7%). Hence, the share of negative correlations of 16.2% is outside any bootstrap confidence interval.

[25] One way to contrast the insignificant overall correlation with the significant within-subject correlation is to correlate guesses larger than zero with the *absolute* value of the difference between the transfer at those guesses and the transfer chosen for a guess equal to zero (i.e., with the absolute change in transfers). If transfers were just chosen randomly, we would expect a zero correlation between guesses and absolute changes in transfers. Yet, this correlation is highly significant with a coefficient of 0.217 (p = 0.004). Performing a Monte-Carlo simulation with different dictator-recipient matching combinations as a robustness check leads to an average correlation coefficient of 0.223 (p = 0.003).

incentivized survey, which the dictators had to complete *after* the conditional transfer decisions have been made but *before* knowing the true guess of their recipient. The survey asked to provide an estimate of the average recipient's guess. It turned out that the estimates were not significantly different from the actual recipient's FOBs (p = 0.315, two-sided MWU-test), with a mean of 4.30 Euro. However, according to the false consensus conjecture, these estimates are expected to be distorted towards the dictator's own transfer choice as they were elicited without giving any information about the recipient. Indeed, the correlation between these *ex ante* self-reported SOBs and the corresponding conditional transfers is highly significant with a coefficient of 0.438 ($p < 0.001$). That is, if transfers were chosen according to the *ex ante* SOB, expectations and transfers would have been strongly correlated. This is similar to the results of previous studies with incentivized SOB elicitation, in which positive correlations between beliefs and actions were found. We also observe that the absolute difference between transfers and SOBs is significantly smaller for the *ex ante* SOB (2.06 on average) compared to the induced SOB (3.20 on average; $p <$ 0.001, two-sided sign test).[26] Overall, we conclude that the false consensus effect is real and may have contributed to the observed significant effect of SOBs in studies based on *ex ante* self-reported SOBs.

3. A MODEL OF INTENTIONAL SURPRISE AND EXPERIMENT 2

The goal of Experiment 1 was to establish that there is heterogeneity regarding how dictators respond to recipients' expectations, as predicted. The goal of Experiment 2 is to show that, although there is heterogeneity in preferences, there are settings in which the incentives are perfectly aligned for all dictators, regardless of whether they care more about negative or positive surprises. At the same time, this section demonstrates that our model can easily be extended to also capture that dictators may care about the attribution of intentionality.

[26] The result is robust to rematching of subjects, as confirmed by Monte-Carlo simulations. The fact that transfers are relatively close to self-reported SOB is neither predicted by, nor inconsistent with our model. However, the false consensus effect does not organize how dictators respond to induced SOBs, which is the focus of our model and experiments.

3.1. Model

3.1.1. Basic idea and specification of utility

We introduce a generalization to the utility function (2) of the dictator in order to capture the possibility that the dictator cares about the recipient's attribution of intentions behind surprises. This corresponds to the intuition from BD who assume, in the analysis of their concept of "guilt from blame", that the recipient blames the first player more if the negative surprise has been intentional (that is, expected by the first player). We then explore the relationship between attribution of intentions and public versus private knowledge about beliefs. In particular, our approach elaborates on the intuition that public knowledge about expectations allows for a straightforward inference about the dictator's intention behind a transfer whereas private information makes this inference difficult. As a result, the dictator's motivation to surprise the recipient, and hence his transfer, is affected by the knowledge condition. The direction of this effect turns out to be independent of how much weight is put on negative and positive surprises, respectively.

Our modified utility function takes the following form:

$$u_i(t_i \mid h_j, h_{jij}) = m_i(1-t_i) + \lambda_1 S_i^S(t_i \mid h_j) + \lambda_2 S_i^I(t_i \mid h_{jij}), \tag{5}$$

where

$$S_i^S(t_i \mid h_j) = \alpha_i \int_0^{t_i} (t_i - x) h_j(x) dx - \beta_i \int_{t_i}^1 (x - t_i) h_j(x) dx, \tag{6}$$

$$S_i^I(t_i \mid h_{jij}) = \alpha_i \int_0^{t_i} (t_i - x) h_{jij}(x \mid t_i) dx - \beta_i \int_{t_i}^1 (x - t_i) h_{jij}(x \mid t_i) dx. \tag{7}$$

Here S_i^S coincides with surprise function S_i considered in the previous section: it denotes utility derived from a *simple* surprise that the dictator experiences directly when deviating from the recipient's expectation. In addition, S_i^I denotes utility derived from the recipient's attribution of the *intentions* behind the surprise, which we refer to below as "intentional surprise".

Technically, the only difference between S_i^I and S_i^S is that in the former case the FOB density $h_j(\cdot)$ is replaced by the posterior third-order belief density $h_{jij}(\cdot \mid t_i)$ conditional on transfer t_i, with corresponding cdf $H_{jij}(\cdot \mid t_i)$. The third-order belief $H_{jij}(x \mid t_i)$ represents the recipient's inference about the actual SOB of the dictator conditional on observing transfer t_i, that is $H_{jij}(x \mid t_i) = E_j[H_{ij}(x) \mid t_i]$. Intuitively, S_i^I thus corresponds to the recipient's inference about whether a deviation from her expectation is by the dictator's intention (as he could foresee the surprise), or due to the dictator's SOB (erroneously) deviating from the recipient's FOB. Analogous to BD's concept of guilt from blame, we assume (through the functional form of (7)) that an *intentional* positive surprise leads to more gratitude from the recipient than a positive surprise that has occurred due to a dictator's confusion about the true expectations of the recipient. Such additional gratitude then leads to a larger utility gain for the dictator from positively surprising the recipient. The same logic applies to negative surprises, where the dictator incurs a higher utility loss if the recipient can infer that her disappointment has been intentional.

The coefficients $\lambda_1 \geq 0$ and $\lambda_2 > 0$ denote the relative weights of S_i^S and S_i^I, respectively, in the dictator's utility, such that their sum is normalized to 1:

$$\lambda_1 + \lambda_2 = 1. \tag{8}$$

We assume λ_2 to be strictly positive because we want to investigate the impact of intentional surprise S_i^I. For the following results we also keep Assumptions A1-A4 laid out in the last section. That is, we assume that the dictator gets a signal θ_j about the median of the recipient's FOB, subject to Assumptions A3 and A4.

3.1.2. Treatment variation and simple surprise

We study two information conditions, which will correspond to our laboratory treatments in Experiment 2: PUBLIC and PRIVATE. In the PRIVATE treatment the recipient does not know for sure that the dictator observes θ_j before his decision. In contrast, in the PUBLIC treatment the signaling of the recipient's *ex ante* FOB is made common knowledge *after* the signal θ_j has been transmitted.[27] Importantly, the dictator knows the information provided to the recipient before he makes his choice.

Since the treatment manipulation occurs *after* the signal θ_j is transmitted, the *ex ante* recipient's FOB, signaled by θ_j, is equivalent in both treatments. Consequently, the simple surprise S_i^S, which incorporates only the *ex ante* recipient's FOB, does not vary between the treatments. That is,

$$S_{i,pub}^S(t_i,\theta_j) = S_{i,pr}^S(t_i,\theta_j)\,, \tag{9}$$

where the lower index *pub* stays for the PUBLIC treatment, and *pr* for the PRIVATE treatment. In contrast, the intentional surprise S_i^I is based on the *ex post* third-order belief, and may thus depend on the treatment manipulation, which in turn, as we show below, leads to a smaller transfer in the PRIVATE treatment.

3.1.3. Intentional surprise in the PUBLIC treatment

In the PUBLIC treatment the transmission of the signal θ_j is common knowledge. Thus, the third-order belief of the recipient in the PUBLIC treatment is, by the law of iterated conditional expectations (see Duffie (1988), p. 84), equal to her FOB:

$$H_{jij,pub}(x,\theta_j \mid t_i) = E_j E_i[H_j(x) \mid \theta_j] = H_j(x,\theta_j)\,. \tag{10}$$

It follows from (6), (7) and (10) that the intentional surprise $S_{i,pub}^I(t_i,\theta_j)$ is equal to the simple surprise $S_{i,pub}^S(t_i,\theta_j)$. That is, the recipient makes a correct inference about the intentions of the dictator:

[27] As it is generally not possible to make sure that some information is common knowledge in the laboratory (because, e.g., somebody may have missed some information), we prefer the term *public* knowledge when we refer to experiment treatments, while we refer to the practically more demanding but theoretically simpler concept of common knowledge in our theory. In our analyses and in our experiment, the important aspect of the PUBLIC treatment is that a dictator knows that his recipient knows that he knows her first-order belief, which is what we explicitly told dictators in our laboratory if they participated in the PUBLIC treatment.

$$S^I_{i,pub}(t_i, \theta_j) = S^S_{i,pub}(t_i, \theta_j)\,. \tag{11}$$

3.1.4. Intentional surprise in the PRIVATE treatment

Because the dictator's SOB is not publicly known in the PRIVATE treatment, the recipient's *ex post* third-order belief about the dictator's SOB can be different from the actual dictator's SOB. We model the formation of the third-order belief in the PRIVATE treatment as follows. Denote by $0 \leq \chi < 1$ the *ex ante* probability assigned by the recipient that her guess θ_j will be transmitted to the dictator in the PRIVATE treatment. Then, given that in the case of belief transmission the recipient's first-order belief becomes common knowledge (implying $H_{jij}(x) = H_{ij}(x) = H_j(x)$), her unconditional third-order belief in the PRIVATE treatment is

$$H_{jij,pr}(x) = \chi H_j(x) + (1-\chi) H^0_{jij}(x), \tag{12}$$

where $H^0_{jij}(\cdot)$ is the expected cdf of the dictator's *ex ante* SOB (with pdf $h^0_{jij}(\cdot)$), which the dictator would have if he did not have any prior information about the recipient's FOB.

As before, we do not directly impose any consistency restrictions between the recipients' and the dictators' *actual* beliefs. However, we impose restrictions on the *internal* consistency of the system of beliefs of a given player. In particular, we assume that the dictator's SOB, as expected by the recipient, is *ex ante* unbiased relative to the recipient's FOB; that is, the recipient believes that the dictator does not systematically under- or overestimates her expectation:

A5. $H^0_{jij}(x) = H_j(x)$.

However, the recipient is uncertain about the *actual* dictator's SOB distribution (which might be heterogeneous in the same way as recipients' FOBs), and hence

$H^0_{jij}(\cdot)$ is represented by a probability weighting over possible dictator's *ex ante* SOBs:[28]

$$H^0_{jij}(x) = \sum_K H^0_{ij,\kappa}(x) p_\kappa \, , \tag{13}$$

where $\kappa \in K$ is a parameter indexing the family of possible *ex ante* SOB distributions, and p_κ is the unconditional probability of $H^0_{ij,\kappa}$ assigned by the recipient. Finally, we assume that the recipient believes that the *ex ante* SOB of the dictator is also internally consistent, i.e., represents a consistent assessment as in Battigalli and Dufwenberg (2009). That is, the recipient believes that each possible *ex ante* SOB distribution $H^0_{ij,\kappa}(\cdot)$ should be unbiased relative to the distribution of transfers conditional on $H^0_{ij,\kappa}(\cdot)$. This can be formulated as:

A6. $H_j(x \mid H^0_{ij,\kappa}) = H^0_{ij,\kappa}(x)$ for any $x \in [0,1]$.

We also make two technical assumptions regarding the dictator's expected SOB distributions:

A7. $h^0_{ij,\kappa}(\cdot)$ is strictly positive and differentiable on $[0,1]$ for any κ.

A8. $h^0_{ij,\kappa}(\cdot)$ can be ordered by some indexing parameter γ so that

$$\frac{d}{dx}\left(\frac{h^0_{ij,\gamma_2}(x)}{h^0_{ij,\gamma_1}(x)}\right) > 0 \text{ if and only if } \gamma_2 > \gamma_1 \text{ for any } x \in [0,1].$$

Assumption A8 is a strict monotone likelihood ratio property (MLRP) and roughly implies that for any two possible SOB distributions one distribution can be said to reflect higher beliefs than the other.

Since the recipient believes that the dictator's SOB affects the distribution of transfers (by Assumption A6), the recipient can infer some information about the

[28] The assumption of a discrete rather than a continuous family of distributions is made for simplicity of exposition and does not limit the generality of the results.

SOB from observing the transfer. Thus, the recipient's third-order belief in the PRIVATE treatment becomes endogenous to the transfer.[29] We employ here an additional technical assumption that if the realized transfer is equal to the expected median transfer, then the recipient infers that her *ex ante* belief is unbiased:

A9. $H_{jij,pr}(x \mid t_i = \theta_j) = H_{jij,pr}(x, \theta_j).$

3.1.5. Results

Our Proposition 2 is mainly based on the following Lemma 1, which implies that the recipient's inference about the dictator's SOB in the PRIVATE treatment is positively correlated with the observed transfer:

Lemma 1. *In the PRIVATE treatment the posterior third-order belief distribution* $H_{jij,pr}(\cdot \mid t_i)$ *is strictly increasing in* t_i *in the sense of FOSD.*

Proof. See Appendix B. ■

Lemma 1 means that after observing a higher transfer, the recipient assigns less probability mass to low SOBs of the dictator. Intuitively, the recipient believes that in case of no belief transmission the equilibrium transfer should be consistent with the dictator's SOB (in line with Assumption A6). Hence, after observing a relatively high transfer in the PRIVATE treatment, the recipient is more likely to attribute the transfer to the dictator's higher SOB. Importantly, the positive relation between the recipient's belief and the transfer holds for both relatively guilt-averse as well as relatively surprise-seeking dictators, and is in both cases driven by the assumption of the anticipated internal consistency of beliefs.

Lemma 1 implies that there is less scope for attribution of intentions in the PRIVATE treatment: the recipient at least partly attributes the observed high (low) transfer to the dictator's own high (low) SOB, rather than to the intention of the dictator to positively (negatively) surprise the recipient. Hence, the dictator's incentives to positively surprise or to avoid guilt, which are based on the intentional

[29] See Appendix B for the formal analysis. For simplicity, we consider the case with two possible *ex ante* SOB distributions $H^0_{ij,\kappa}(\cdot)$.

part of the surprise function, are reduced, resulting in a generally lower transfer in the PRIVATE treatment:

Proposition 2. *If the optimal transfer in the PUBLIC treatment is interior, then it is strictly larger than that in the PRIVATE treatment*.

Proof. See Appendix B. ■

Proposition 2 holds independently of to what extent dictators care for negative and positive surprises, and, in particular, of the sign of $\alpha_i - \beta_i$. Also, here too, behavior cannot be confounded by dictator false consensus effects, because the dictator is informed of the recipient's FOB in both treatments.[30] Thus, introducing attribution of intentions into the analysis allows for a clear and robust testable prediction of the role of others' expectations and intentions at the between-subject level, complementing the within-subject level analysis in Experiment 1.

3.2. Experiment design, hypotheses and results

3.2.1. Design and hypotheses

Our second experiment was conducted with 254 participants in the Cologne Laboratory for Economic Research and the Frankfurt Laboratory of Experimental Economics. As in EJTT's original experiment, our recipients were asked about their expectation regarding the average amount a dictator would send. Again, the recipient whose estimate was closest to the true average amount sent received an extra payment of 8 Euro. Each dictator was then informed about the expectation of the recipient matched to him before a decision on how to split 10 Euro is made. Participants were recruited with the help of the online recruitment system ORSEE (Greiner (2004)); the experiment program was developed with the software z-Tree (Fischbacher (2007)).

In line with our model, we conducted two treatments with a between-subject design. In the PUBLIC treatment, recipients were told, after the expectation was elicited, that the matched dictator would get to know their estimate before choosing

[30] However, because false consensus may matter for recipients' inference in PRIVATE, we note that Proposition 2 would be robust to such false consensus as we show in Appendix C.

the transfer (i.e., χ was set to 1). In the PRIVATE treatment, recipients were not informed that their estimations were communicated to dictators (i.e., χ was set to 0).[31] The respective procedure was known to dictators. Instructions can be found in Appendix E.

The design allows us to change the scope for the recipients' inferences on dictator intentions, while at the same time minimizing strategic reporting of expectations and keeping recipients' expectations fixed. Contrary to Experiment 1, where we elicited within-subject correlations between transfers and beliefs and thus used the strategy method, we conducted this experiment with the direct response method to focus on the treatment effect.

Our null hypothesis is that dictators are indifferent to recipients' inferences regarding the underlying intentions. Then, there should be no difference in transfers between the treatments. In contrast, Proposition 2 predicts that transfers are higher in the PUBLIC treatment. In fact, because our treatment variation exclusively and directly affects third-order beliefs of the recipient, and the dictator's inference about these beliefs, a treatment effect would demonstrate the relevance of these higher-order beliefs, which so far has not been established in the literature.

3.2.2. Results

The average amount sent in PUBLIC was with 1.68 Euro almost 70% higher than the 1.01 Euro observed in the PRIVATE treatment (p = 0.022, two-sided MWU-test).[32] This confirms the prediction of Proposition 2. Also, in line with the reduced incentive to take into account the expectation of the recipient in the PRIVATE treatment, we find that 71.7% (43 out of 60) of the dictators in PRIVATE transferred less than the recipient's expectation compared to 50.7% of the dictators in PUBLIC (34 out of 67; p = 0.016, two-sided χ^2-test). Similarly, and importantly for our model of surprising gifts, the share of dictators who exceeded the recipient's expectation is

[31] If some recipients suspected that their beliefs could be transmitted (so that $0 < \chi < 1$), the prediction of Proposition 2 is still valid.

[32] When we compute transfers as a percentage of the total amount to be divided (the cake size differed across the experiments), we find that the average amounts sent were higher in Experiment 1 than in Experiment 2 (23.2% versus 13.6%; p < 0.001, two-sided MWU-test). Part of the reason might be the larger social distance between dictators and recipients in the laboratory in Experiment 2, compared to the classroom Experiment 1.

more than twice as high in the PUBLIC than in the PRIVATE treatment: 28.4% versus 11.7% (p = 0.020, two-sided χ^2-test).

Due to the heterogeneity of preferences, we should – similar to Experiment 1 and the related studies on guilt aversion – observe only little correlation between recipients' beliefs and transfers in the aggregate in both treatments. Indeed, this is what we find: Pearson correlation coefficients are 0.149 (p = 0.230) in the PUBLIC treatment and 0.020 (p = 0.881) in the PRIVATE treatment. All results are corroborated by Tobit models with the amount sent by the dictator as the dependent variable and including a number of demographic variables of the subjects (see Table F.1 in Appendix F).

We conclude that dictators do not only respond to recipients' expectations in ways that are consistent with our model, but also care about what recipients know about intentions behind surprises.

4. ROBUSTNESS CHECKS

Our experiment designs, following EJTT's original design, might be seen as controversial because strategically relevant information is withheld from recipients. Also, the design introduces information asymmetries among participants, as dictators know more about the relevant decision situation than recipients. Thus, as a robustness check, we conducted two further experiments (Experiments 3 and 4), which partly mitigate the problems associated with EJTT's original design. Specifically, dictators in the new experiments have no private knowledge about the strategically relevant aspects of the game at any experiment stage. However, we caution that our new experiments still omit strategically relevant information. The experiments had two phases, which was commonly known. Recipients submitted their guesses in the first phase of the experiment, yet were not told at that time that they will have the option to transmit these beliefs in the second phase of the experiment (otherwise they might have had an incentive to strategically inflate guesses). We acknowledge that this design feature could still be considered problematic, because subjects in future experiments might be concerned about such unanticipated relevance of judgments about others' behavior – although the option to not transmit one's guess provides recipients with full control over transmission of their beliefs and thus a form of insurance.[33]

Both additional experiments were conducted in the Cologne Laboratory for Economic Research with altogether 306 participants (180 participants in Experiment 3, 126 participants in Experiment 4). Participants were recruited with the online recruitment system ORSEE (Greiner (2004)). The experiment software was programmed with z-Tree (Fischbacher (2007)). The following subsections describe the experiment procedures and results in detail.

4.1. Experiment 3: Between-subjects heterogeneity

4.1.1. Design and hypotheses

The main difference between Experiment 3 and Experiment 1 is that we now split the decision situation into two parts. In the first part, two participants were randomly matched and played a simple dictator game in which 14 Euro had to be allocated between the dictator and the recipient. Prior to the dictator decision, we elicited the guess of each recipient about the average dictator transfer, using the same procedures as in Experiment 1. Importantly, a dictator in Part 1 never observed the guess of his matched recipient, and the recipient received feedback about the transfer of this dictator only at the end of the experiment. Also, subjects did not receive instructions for Part 2 of the experiment until Part 1 has been finished; they were only notified that an additional part with new instructions will follow.

In Part 2 of the experiment, dictators and recipients were randomly rematched, ensuring that no one would interact with a participant from Part 1. Contrary to Experiment 1, in which the dictator automatically received the information about the recipient's guess, recipients now had to *choose* at the beginning of Part 2 whether or not to transmit their Part 1-guesses to their Part 2-dictators. If the recipient chose not to transmit, the dictator was informed about this and then had to state an unconditional transfer (from 0 to 14 Euro). However, if the recipient agreed to transmit her guess, the decision of the dictator was elicited in the same way as in

[33] As we will see in the next subsections, only a negligible share of recipients opted against the transmission of beliefs in Experiments 3 and 4.

Experiment 1. That is, the dictator had to state transfers conditional on all possible guesses of the recipient (rounded to values of 50 Euro cents), with guesses higher than 9 Euro being combined in a single category.[34] Hence, as before, we collected 19 decisions for these dictators. The dictator was informed about the actual guess of the recipient at the end of the experiment. The transfer that corresponded to the actual rounded guess of the matched recipient was then relevant for the final payoffs. Importantly, both recipients and dictators received complete information about all procedures at the start of Part 2, including the fact that dictators would state their transfers conditional on guesses. At the end of Experiment 3, participants were informed about the results of the two parts and whether or not they received the bonus for their guess. One of the two parts was then randomly chosen for payoffs.[35]

Our hypothesis is, based on Proposition 1, that our findings from Experiment 1 can be replicated with our novel design. In particular, we predict that transfers are (positively or negatively) correlated with recipients' expectations.

4.1.2. Results

The large majority of recipients (88 out of our 90 subjects or 97.8%) agreed to transmit their guesses to the dictators in Experiment 3. As a result, we collected conditional transfers for 88 dictators, yielding 1,672 observations.

Despite the differences in the design of Experiments 1 and 3, subjects behaved in a remarkably similar way. Here we report the most important results; see Table F.2 in Appendix F for a table reporting all data comparisons. On average, in the relevant Part 2 of Experiment 3, dictators transferred 2.24 Euro conditioned on the true FOB

[34] Because the recipient can choose whether or not to transmit the belief before the dictator decides about the transfer, the game is dynamic and the decision to transmit or not might involve selection and signaling effects (Battigalli and Dufwenberg (2009)). Our model is static and thus cannot speak to this issue. However, if no or only a negligible number of subjects decides to not allow belief transmission, our model describes the subgame following the recipient's choice accurately. Thus, we paid recipients who gave the permission to transmit their guess in Part 2 an additional 2.50 Euro. (Dictators were paid an additional 2.50 Euro in any case, making sure that the additional payment does not introduce another asymmetry between dictators and guess-transmitting recipients.) As the vast majority of recipients agreed to the transmission of guesses (see below), we can neglect the dynamic play in our data analyses.

[35] In Experiment 3, we additionally elicited the dictators' guesses about the recipients' guesses in Part 1, providing the same incentives as for recipients' guesses. Also, the bonus for the best guesses was paid out irrespective of which part was chosen in order to balance expected payoffs from guesses across parts. See the Instructions for Experiment 3 in Appendix E.

of their matched recipient. As in Experiment 1, we do not observe a significant correlation between realized transfers and matched beliefs, with the Pearson correlation coefficient again close to zero (0.054, p = 0.619). Also, we find very similar shares of dictators across the experiments who varied their transfer in response to guesses at least once (71.6% in Experiment 3 and 77.5% in Experiment 1) and who exhibited a significant within-subject correlation of transfers and guesses (58.0% in Experiment 3 and 53.9% in Experiment 1). The share of subjects with a positive correlation (relatively guilt-averse) among all dictators accounts for 43.2% (37.7% in Experiment 1), whereas the share of surprise-seeking dictators with a negative within-subject correlation is 14.8% (16.2% in Experiment 1). Similarly, the share of transfers that strictly exceeded guesses accounts for 23.9% in Experiment 3 (27.2% in Experiment 1).

There is a difference in the average sizes of positive and negative regression coefficients across experiments suggesting that the strength of these relations is somewhat smaller in Experiment 3: positive coefficients are on average 0.36 (0.58 in Experiment 1) and negative coefficients are on average 0.29 (0.53 in Experiment 1). Yet, despite the somewhat weaker effect of the relation between transfers and guesses, dictators in Experiment 3 tended to behave in a more consistent manner than in Experiment 1: only 15.7% of the dictators with significant within-subject correlation changed transfers in response to increasing beliefs in a non-monotonic way, whereas the corresponding share in Experiment 1 is 30.1%.

Overall, the results of Experiments 3 and 1 are highly consistent with each other. Therefore, Experiment 3 strongly confirms the importance of considering between-subjects heterogeneity in response to recipients' expectations in order to detect the relevance of guilt aversion and surprise seeking for dictator decisions – and it does so in a novel experiment design that mitigates the problems associated with EJTT's approach.

4.2. Experiment 4: Intentions behind surprises

4.2.1. Design and hypotheses

Experiment 4 serves as a robustness check for Experiment 2 and had two parts, too. The first part was identical to Part 1 of Experiment 3, again with the purpose to elicit unconfounded recipient guesses about average dictator transfers. Also like in Experiment 3, in Part 2 dictators received an endowment of 14 Euro that they could allocate between themselves and the newly matched recipients, and recipients had to decide whether or not to allow the transmission of guesses elicited in Part 1 to the newly matched dictators.[36] The main new feature of Experiment 4 is that even when the recipient agreed to the transmission of her Part 1-guess in Part 2, the Part 2-dictator was informed about the guess only with a probability of 50%; otherwise the guess was not transmitted. This created an *ex ante* uncertainty to the recipient about whether or not the dictator would actually see her guess.[37]

If both recipient and nature had transmitted the guess, we varied the recipient's *ex post* knowledge about the dictator's SOB (akin in Experiment 2). In the PUBLIC treatment, the recipient would get to know at the end of the experiment if the dictator had seen her FOB prior to his decision. On the other hand, in the PRIVATE treatment, the recipient would stay ignorant about whether or not her belief had actually been transmitted. Each dictator could provide two conditional transfers, depending on the payoff relevant treatment variation (PUBLIC or PRIVATE), which was determined by a fair chance move. This way, we were able to investigate within-subject treatment effects.[38] As in Experiment 3, dictators and recipients were

---

[36] As before, if recipients opted for the transmission, they would receive an additional payment of 2.50 Euro irrespective of dictator decisions. Dictators were paid the 2.50 Euro in any case.

[37] Subjects were told that they are playing in two "rounds" in Part 2, each round with a different opponent, and that the probability of belief transmission is 50% in each round. The 50% probability was implemented in a way that the guess of a matched recipient was transmitted to the dictator in one of the two rounds (conditional on the corresponding recipient's agreement to transmit beliefs, which was elicited only once and applied to both dictators she would be matched with). No feedback was provided between the rounds. See the Instructions in Appendix E for the details.

[38] Thus, our Experiment 4 provides an additional check of the robustness of the treatment effect observed in Experiment 2 with a direct response method. The potential effect of the strategy method on transfers is *ex ante* hard to predict, though. On the one hand, the strategy method increases the saliency of the experimental variation and therefore might emphasize differences between treatments. On the

informed about all aspects of the decision situation before Part 2 started, and feedback about transfers in both parts and about actual belief transmission in Part 2 was given to the recipient only at the end of the experiment. See the Instructions in Appendix E for procedural details and Appendix D for an illustration of the sequence of actions.

Our central hypothesis, based on Proposition 2, is the same as in Experiment 2: Transfers in the PUBLIC treatment should be higher than transfers in the PRIVATE treatment for dictators who care about beliefs, as both guilt-averse and surprise-seeking subjects have less incentive to transfer their endowment to the recipient if dictators' SOBs remain private knowledge. Moreover, we predict that the effect tends to be smaller than what is suggested by Experiment 2. The reason is that the treatment effect is decreasing in recipient's *ex ante* probability χ of belief transmission in the PRIVATE treatment, which is at its minimum of 0 in Experiment 2 and 50% in Experiment 4. The underlying mechanism is that the recipient's attribution of the transfer to the dictator's true intention becomes stronger with larger χ.[39]

4.2.2. Results

Similar to Experiment 3, we do not have to deal with selection effects, as all recipients (63 out of 63) agreed to transmit their beliefs to the dictators. Our main hypothesis is corroborated by the data: On average, dictators sent 3.05 Euro (21.8% of the endowment) to the recipient in the PUBLIC treatment and 2.63 Euro (18.8% of the endowment) in the PRIVATE treatment, and this difference is significant ($p < 0.001$, two sided Wilcoxon Matched-Pairs Signed-Ranks test). 54.0% of the dictators did not respond at all to the information condition, which would be consistent with λ_2 being close to zero. Among those dictators who did respond, 86.2% transferred a higher amount in PUBLIC than in PRIVATE.

Regarding the size of the effect, we observe that while the treatment effect is highly significant, it is relatively small: in the full sample, subjects transferred on average 16% more in the PUBLIC treatment, compared to the 70% increase between

other hand, subjects may prefer to behave in a "consistent" way, not exploiting asymmetric information when it comes to social behavior. See also the discussion in our concluding section.

[39] See (B.6) in Appendix B; the dictator is less able to manipulate the recipient's inference about his SOB by changing his transfer if χ is larger.

the treatments in Experiment 2. This is consistent with our prediction, although we caution that the data are only suggestive here, as some design parameters are changed beyond χ.

Similar to what we see in all other experiments, in Experiment 4, too, the role of FOBs is limited in the aggregate. There is a small, insignificantly positive correlation between recipients' expectations and transfers in PUBLIC that is similar in size as in Experiment 2 (the Pearson correlation coefficient is 0.211, p = 0.097), and there is no correlation in PRIVATE (the Pearson correlation coefficient is 0.043, p = 0.739). Table F.3 in Appendix F shows Tobit models that support all conclusions.

To sum up, our robustness checks in Experiments 3 and 4 replicate all important patterns from our Experiments 1 and 2; the data show remarkably robust regularities in line with our central hypotheses from Propositions 1 and 2.

5. DISCUSSION AND CONCLUSION

We propose a model of 'surprising gifts' to investigate the role of others' expectations for giving in dictator games. The model assumes that people care not only about negative but also about positive surprises induced by their actions. We find strong evidence for our model in a series of experiments, employing more than 900 participants. While, similar to EJTT, we do not find a correlation between induced SOBs and actual transfers on a between-subject level, a within-subject level analysis in the first experiment shows that a large fraction of dictators reacts to recipients' expectations. In particular, many dictators behave consistently with BD's notion of guilt aversion. Yet, there is also a significant share of dictators behaving consistently with a preference for exceeding others' expectations. The heterogeneity of belief-dependent preferences among subjects explains the lack of correlation between SOBs and transfers in the aggregate.

We then extend our model to integrate the notion that dictators may care for the recipients' inferences about the intention behind a transfer. The model predicts smaller transfers for both relatively guilt-averse and relatively surprise-seeking dictators if the inference about the dictator's intentions becomes ambiguous. Our data from the second experiment confirm the prediction and, additionally, show that transfers are at least partly driven by an attribution effect.

We also extend EJTT's clever but controversial design in Experiments 3 and 4 in order to test whether our (and EJTT's) results replicate if the asymmetry of strategically relevant information between dictators and recipients is avoided and recipients have the possibility to veto the belief transmission. We find that all main conclusions remain valid. We caution, however, that our design still involves omission of strategically relevant information in one phase of the experiment. Future research might propose experiment methods to elicit beliefs in an unbiased way avoiding such omissions.

One might hypothesize that some of our experimental findings are affected by the strategy method: subjects may consider it 'appropriate' to take others' expectations into account, just because they are given an explicit choice to condition their decisions on expectations. While there generally appears to be no obvious a priori reason to favor one or the other method (Brandts and Charness (2011)), it does not seem implausible that the elicitation method can matter. We doubt, however, that the elicitation method is critical for our main conclusions. For one, we have shown that our strategy method results are, to the extent comparable, fully consistent with EJTT's dictator game results, where the direct-response method is employed. Moreover, our data from our Experiment 2 provide complementary evidence for our model with the direct-response approach, and the data are fully consistent with our data from the analogous strategy-method Experiment 4 (see Brandts and Charness (2011) and Fischbacher et al. (2012) for similar findings in other contexts).

Overall, our data are consistent with the hypothesis that guilt aversion is a major motivation for giving in dictator games. At the same time, our analysis highlights that many subjects also like to exceed others' expectations, and that taking this motive into account, along with a motivation that subjects care about the attribution of their intentions, may improve the predictive value of the model. A natural next step would be to investigate how our results generalize to other social contexts. For example, related papers have applied guilt aversion to hidden action problems in principal-agent relationships (e.g., Charness and Dufwenberg (2006)), and to strategic communication (Beck et al. (2013)).

It seems also worthwhile to investigate how a concern to please the opponent interacts with preferences to adhere to general norms of social behavior (Bernheim (1994), Sliwka (2007)). In our experimental context one might argue, for instance, that the recipient's belief, if transmitted to the dictator, may be a valuable signal of what is the generally acceptable transfer. If dictators are also driven by a preference to conform to such a generally accepted social norm – as opposed to a desire to please one's own, particular recipient – dictators have even more reason to condition their social behavior on beliefs. We note, however, that a norm-based explanation alone is difficult to reconcile with the negative correlation of induced SOBs and transfers that we observe for many dictators in Experiment 1. More generally, for the explanations based on social norms to go through, all that should matter is the knowledge about the recipient's expectation, while our experiments show that, keeping expectations constant, dictators care about the recipient's attribution of their intentions to surprise, too. Clearly, belief-dependent preferences matter for gift-giving.

ACKNOWLEDGEMENTS

Financial support of the German Research Foundation (DFG) through the Leibniz program and through the Research Unit "Design & Behavior" (FOR 1371) is gratefully acknowledged. We thank Martin Dufwenberg, Roman Inderst, Jean-Robert Tyran, the Editor, three anonymous referees, and conference and seminar participants in Maastricht, Cologne and Stanford for helpful comments and suggestions. Ockenfels thanks the Economics Department at Stanford University for their generous hospitality.

REFERENCES

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

G. Attanasi, P. Battigalli, R. Nagel, Disclosure of Belief–Dependent Preferences in a Trust Game, Working Paper, 2014.

P. Battigalli, M. Dufwenberg, Guilt in Games, Amer. Econ. Rev., 97 (2007) 170–176.

P. Battigalli, M. Dufwenberg, Dynamic Psychological Games, J. Econ. Theory, 141 (2009) 1–35.

A. Beck, R. Kerschbamer, J. Qiu, M. Sutter, Shaping beliefs in experimental markets for expert services: Guilt aversion and the impact of promises and money-burning options, Games Econ. Behav., 81 (2013) 145–164.

C. Bellemare, A. Sebald, M. Strobel, Measuring the Willingness to Pay to Avoid Guilt: Estimation Using Equilibrium and Stated Belief Models, J. Appl. Econ., 26 (2011) 437–453.

C. Bellemare, A. Sebald, S. Suetens, Heterogeneous guilt aversion and incentive effects, Working Paper, 2014.

R. Bénabou, J. Tirole, Incentives and Prosocial Behavior, Amer. Econ. Rev., 96 (2006) 1652–1678.

D. B. Bernheim, A Theory of Conformity, J. of Polit. Economy, 100 (1994) 841–877.

G. E. Bolton, A. Ockenfels, ERC: A Theory of Equity, Reciprocity, and Competition, Amer. Econ. Rev., 90 (2000) 166–193.

J. Brandts, G. Charness, The Strategy Versus the Direct-Response Method: A First Survey of Experimental Comparisons, Exp. Econ., 14 (2011) 375–398.

A.W. Cappelen, T. Halvorsen, E. Sorensen, B. Tungodden, Face-Saving or Fair-Minded: What Motivates Moral Behavior?, Working Paper, Norwegian School of Economics, 2013.

G. Charness, M. Dufwenberg, Promises and Partnership, Econometrica, 74 (2006) 1579–1601.

G. Charness, M. Dufwenberg, Participation, Amer. Econ. Rev., 101 (2011) 1211–1237.

J. Dana, R.A. Weber, J.A. Kuang, Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness, Econ. Theory, 33 (2007) 67–80.

D. Duffie, Security Markets: Stochastic Models, Academic Press, Boston, 1988.

M. Dufwenberg, U. Gneezy, Measuring Beliefs in an Experimental Lost Wallet Game, Games Econ. Behav., 30 (2000) 163–182.

M. Dufwenberg, S. Gächter, H. Hennig-Schmidt, The Framing of Games and the Psychology of Play, Games Econ. Behav., 73 (2011) 459–478.

T. Ellingsen, M. Johannesson, S. Tjøtta, G. Torsvik, Testing Guilt Aversion, Games Econ. Behav., 68 (2010) 95–107.

A. Falk, M. Kosfeld, The Hidden Costs of Control, Amer. Econ. Rev., 96 (2006) 1611–1630.

E. Fehr, K.M. Schmidt, A Theory of Fairness, Competition and Cooperation, Q. J. Econ., 114 (1999) 817–868.

U. Fischbacher, z-Tree: Zurich toolbox for ready-made economic experiments, Exp. Econ., 10 (2007) 171−178.

U. Fischbacher, S. Gächter, S. Quercia, The Behavioral Validity of the Strategy Method in Public Good Experiments, J. of Econ. Psychol., 33 (2012) 897–913.

J. Geanakoplos, D. Pearce, E. Stacchetti, Psychological Games and Sequential Rationality, Games Econ. Behav., 1 (1989) 60–79.

B. Greiner, An Online Recruitment System for Economic Experiments, in: K. Kremer, V. Macho (Eds.), Forschung und wissenschaftliches Rechnen GWDG Bericht Nr. 63, Gesellschaft für Wissenschaftliche Datenverarbeitung mbh Göttingen, Göttingen, 2004, pp. 79−93.

Z. Grossman, Self-Signaling Versus Social Signaling in Giving, Working Paper, UCSB, 2010.

Z. Grossman, Strategic Ignorance and the Robustness of Social Preferences, Manag. Sci., 60 (2014) 2659–2665.

G. Guerra, D.J. Zizzo, Trust Responsiveness and Beliefs, J. Econ. Behav. Organ., 55 (2004) 25–30.

W. Güth, S. Huck, From Ultimatum Bargaining to Dictatorship – An Experimental Study of Four Games Varying in Veto Power, Metroeconomica, 48 (1997) 262−279.

W. Güth, S. Huck, P. Ockenfels, Two-Level Ultimatum Bargaining With Incomplete Information: An Experimental Study, Econ. J., 106 (1996) 593–604.

D. Kahneman, A. Tversky, Prospect Theory: An Analysis of Decision under Risk, Econometrica, 47 (1979) 263–291.

T. Kawagoe, Y. Narita, Guilt aversion revisited: An experimental test of a new model, J. Econ. Behav. Organ., 102 (2014) 1–9.

B. Köszegi, M. Rabin, A Model of Reference-Dependent Preferences, Q. J. Econ., 121 (2006) 1133–1165.

J.T. Larsen, A. P. McGraw, B. A. Mellers, J. T. Cacioppo, The Agony of Victory and Thrill of Defeat: Mixed Emotional Reactions to Disappointing Wins and Relieving Losses, Psychol. Sci., 15 (2004) 325–330.

B.A. Mellers, A. Schwartz, K. Ho, I. Ritov, Decision Affect Theory: Emotional Reactions to the Outcomes of Risky Options, Psychol. Sci., 8 (1997) 423–429.

M. Mitzkewitz, R. Nagel, Experimental Results on Ultimatum Games with Incomplete Information, Int. J. Game Theory, 22 (1993) 171–198.

A. Ockenfels, D. Sliwka, P. Werner, Bonus Payments and Reference Point Violations, Manag. Sci., forthcoming.

A. Ockenfels, P. Werner, Hiding Behind a Small Cake in a Newspaper Dictator Game, J. Econ. Behav. Organ., 82 (2012) 82–85.

E. Reuben, P. Sapienza, L. Zingales, Is mistrust self-fulfilling?, Econ. Letters, 104 (2009) 89–91.

L. Ross, D. Greene, P. House, The "False Consensus Effect": An Egoistic Bias in Social Perception and Attribution Processes, J. Exp. Soc. Psychol., 13 (1977) 279–301.

B.J. Ruffle, Gift Giving with Emotions, J. Econ. Behav. Organ., 39 (1999) 399–420.

R. Selten, Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments, in: H. Sauerman (Ed.), Beiträge zur experimentellen Wirtschaftsforschung, Mohr Siebeck GmbH & Co. KG, Tübingen, 1967, pp. 136–168.

R. Selten, A. Ockenfels, An Experimental Solidarity Game, J. Econ. Behav. Organ., 34 (1998) 517–539.

D. Sliwka, Trust as a Signal of a Social Norm and the Hidden Costs of Incentive Schemes, Amer. Econ. Rev., 97 (2007) 999–1012.

D. Taubinsky, Great Expectations and Prosocial Acts: Theory and Experiments, Working paper, Harvard University, 2012.

A. Tversky, D. Kahneman, Advances in Prospect Theory: Cumulative Representation of Uncertainty, J. Risk Uncertain., 5 (1992) 297–323.
C. Vanberg, Why do People Keep their Promises? An Experimental Test of Two Explanations, Econometrica, 76 (2008) 1467–1480.

APPENDIX A. Omitted proofs: Experiment 1.

Lemma 2. *For any* α_i *and* β_i *there exists a concave monetary function* m_i, *satisfying assumption A2, and* $\theta_j \in [0,1]$, *such that for any* $\hat{t} \in \arg\max_{t_i} U(t_i, \theta_j)$ *it holds* $0 < \hat{t} < 1$.

Proof. A necessary and sufficient condition for the claim of the lemma is that there exists at least one $0 < t' < 1$ such that both of the following inequalities hold:

$$U_i(t', \theta_j) - U_i(0, \theta_j) > 0, \tag{A.1}$$

$$U_i(t', \theta_j) - U_i(1, \theta_j) > 0. \tag{A.2}$$

Let us consider the first inequality, which ensures that $t_i = 0$ is a suboptimal transfer. Simplifying the expected utility function using integration by parts we get

$$\begin{aligned} U_i(t_i, \theta_j) &= m_i(1-t_i) + \alpha_i \int_0^{t_i} (t_i - x) h_{ij}(x) dx - \beta_i \int_{t_i}^1 (x - t_i) h_{ij}(x) dx \\ &= m_i(1-t_i) + \alpha_i \int_0^{t_i} H_{ij}(x \mid \theta_j) dx + \beta_i \int_{t_i}^1 H_{ij}(x \mid \theta_j) dx - \beta_i \left(1 - t_i\right). \end{aligned} \tag{A.3}$$

Substituting this expression for the utility functions in (A.1) results in

$$(\alpha_i - \beta_i) \int_0^{t'} H_{ij}(x \mid \theta_j) dx + \beta_i t' > m_i(1) - m_i(1-t'). \tag{A.4}$$

Let us define

$$\sigma_i = \int_{1-t'}^{1} \left(m_i'(x) - \frac{\alpha_i + \beta_i}{2} \right) dx, \tag{A.5}$$

which denotes the accumulated difference between the value of the marginal monetary utility and its lower bound (according to Assumption A2). Then,

$$m_i(1) - m_i(1-t') = \sigma_i + \frac{\alpha_i + \beta_i}{2} t'. \tag{A.6}$$

Substituting (A.6) into (A.4) we obtain

$$(\alpha_i - \beta_i) \int_0^{t'} \left(H_{ij}(x \mid \theta_j) - \frac{1}{2} \right) dx > \sigma_i. \tag{A.7}$$

Let us consider the LHS of the inequality. Note that since θ_j is the median of the recipient's FOB distribution, it is the median of the conditional SOB distribution as well:

$$H_{ij}(\theta_j \mid \theta_j) = E_i[H_j(\theta_j) \mid \theta_j] = E_i[\frac{1}{2}] = \frac{1}{2}. \tag{A.8}$$

Consequently, it holds for any t':

$$\lim_{\theta_j \to 0} \int_0^{t'} \left(H_{ij}(x \mid \theta_j) - \frac{1}{2} \right) dx > 0, \tag{A.9}$$

$$\lim_{\theta_j \to 1} \int_0^{t'} \left(H_{ij}(x \mid \theta_j) - \frac{1}{2} \right) dx < 0. \tag{A.10}$$

These inequalities ensure that for any α_i, β_i and t' there exists θ_j such that the LHS of (A.7) is strictly positive. Consequently, for any α_i, β_i and t', if σ_i is sufficiently small, then (A.7) is satisfied at least for some θ_j.

Let us consider the second inequality (A.2), which ensures that $t_i = 1$ is a suboptimal transfer. Substituting (A.3) into (A.2) we get

$$(\alpha_i - \beta_i)\int_{t'}^{1} H_{ij}(x \mid \theta_j)dx + \beta_i(1-t') < m_i(1-t'). \tag{A.11}$$

Denote

$$\varpi_i = \int_0^{1-t'} \left(m_i'(x) - \frac{\alpha_i + \beta_i}{2} \right) dx, \tag{A.12}$$

so that

$$m_i(1-t') = \varpi_i + \frac{\alpha_i + \beta_i}{2}(1-t'). \tag{A.13}$$

Substituting (A.13) into (A.11) we obtain

$$(\alpha_i - \beta_i)\int_{t'}^{1} (H_{ij}(x \mid \theta_j) - 1/2)dx < \varpi_i. \tag{A.14}$$

Since the LHS of (A.14) is bounded for given α_i, β_i and t', there always exists a function $m_i(\cdot)$ with sufficiently large ϖ_i so that (A.14) is satisfied for any θ_j. Thus, if for given α_i, β_i and t' a function $m_i(\cdot)$ is characterized by a sufficiently small positive σ_i (so that (A.7) is satisfied for some θ_j) and a sufficiently large ϖ_i (so that (A.14) is satisfied for the same θ_j), then there exists an interior solution for these θ_j. Since Assumption A2 and concavity of $m_i(\cdot)$ allow for both infinitely small σ_i and (simultaneously) infinitely large ϖ_i for given α_i, β_i and t', the claim holds. ■

Lemma 3. *$U_i(t_i, \theta_j)$ has the strict single crossing property*[40] *in (t_i, θ_j) ($(t_i, -\theta_j)$) if $\alpha_i < \beta_i$ ($\alpha_i > \beta_i$).*

Proof. We have

[40] Function $f(x,z)$ has the strict single crossing property in (x,z) if for any $x' > x''$ and $z' > z''$ it holds that $f(x',z'') - f(x'',z'') \geq 0$ implies $f(x',z') - f(x'',z') > 0$ (Milgrom and Shannon (1994)).

$$\frac{\partial U_i}{\partial t_i}(t_i,\theta_j) = -m_i'(1-t_i) + \alpha_i \int_0^{t_i} h_{ij}(x|\theta_j)dx + \beta_i \int_{t_i}^{1} h_{ij}(x|\theta_j)dx$$
$$= (\alpha_i - \beta_i)H_{ij}(t_i|\theta_j) + \beta_i - m_i'(1-t_i). \quad \text{(A.15)}$$

Taking the partial derivative of this expression with respect to θ_j we get[41]

$$\frac{\partial^2 U_i(t_i,\theta_j)}{\partial t_i \partial \theta_j} = (\alpha_i - \beta_i)\frac{\partial H_{ij}(t_i|\theta_j)}{\partial \theta_j}. \quad \text{(A.16)}$$

This, together with Assumption A3, implies that

$$\text{sgn}\left(\frac{\partial^2 U_i(t_i,\theta_j)}{\partial t_i \partial \theta_j}\right) = -\text{sgn}(\alpha_i - \beta_i) \quad \text{(A.17)}$$

for any $t_i \in (0,1),$ which leads to the claim. ■

Proof of Proposition 1. Let us consider arbitrary values θ_j' and θ_j'' such that $\theta_j' > \theta_j''.$ By Lemma 3 and Milgrom-Shannon (1994) Monotone Selection Theorem we have that for any $t_i' \in \arg\max_{t_i} U_i(t_i,\theta_j')$ and $t_i'' \in \arg\max_{t_i} U_i(t_i,\theta_j'')$ it holds $t_i' \geq t_i''$ $(t_i' \leq t_i'')$ if $\alpha_i < \beta_i$ $(\alpha_i > \beta_i).$ That is, since $t_i^*(\cdot)$ is a best response function such that $t_i^*(\theta_j) \in \arg\max_{t_i} U_i(t_i,\theta_j),$ it is weakly increasing (decreasing) in θ_j if $\alpha_i < \beta_i$ $(\alpha_i > \beta_i)$. If, in addition, $0 < t_i'' < 1$ (the consistency of this condition with our assumptions is established by Lemma 2), then the FOC for maximizing $U_i(t_i,\theta_j)$ must be satisfied at (t_i'',θ_i''):

$$\frac{\partial U_i}{\partial t_i}(t_i'',\theta_i'') = 0. \quad \text{(A.18)}$$

Equation (A.18), together with (A.17), implies that if $\alpha_i \neq \beta_i$, then

[41] The existence of the derivative on the right hand side of (A.16) is ensured by Assumption A4.

$$\frac{\partial U_i}{\partial t_i}(t_i'',\theta_j') \neq 0. \quad \text{(A.19)}$$

It follows that $t_i'' \notin \arg\max_{t_i} U_i(t_i,\theta_j')$, hence $t_i'' \neq t_i'$ (if $\alpha_i \neq \beta_i$).[42] This, together with the previously established fact that $t_i' \leq t_i''$ ($t_i' \geq t_i''$) if $\alpha_i < \beta_i$ ($\alpha_i > \beta_i$), yields that $t_i^*(\theta_j)$ is *strictly* increasing (decreasing) in θ_j if $\alpha_i < \beta_i$ ($\alpha_i > \beta_i$) and $0 < t_i^*(\theta_j) < 1$.
■

REFERENCES

A.S. Edlin, C. Shannon, Strict Monotonicity in Comparative Statics, J. Econ. Theory, 81 (1998) 201–219.

P. Milgrom, C. Shannon, Monotone comparative statics, Econometrica, 64 (1994) 157–181.

APPENDIX B. Omitted proofs: Experiment 2.

Let us simplify the subsequent notation so that the pdf of belief of order k is denoted as $h_k(\cdot)$ and the cdf as $H_k(\cdot)$. The expected utility function of the dictator is denoted by $U_i(\cdot)$. For simplicity, we consider the case with two *ex ante* dictator SOB distributions $H^0_{ij,\kappa}(\cdot)$ characterized by the cdfs $H^0_{21}(\cdot)$ and $H^0_{22}(\cdot)$ (and pdfs $h^0_{21}(\cdot)$ and $h^0_{22}(\cdot)$), which have *ex ante* probabilities p_1 and $p_2 = 1 - p_1$, respectively. In this case Assumption A8 translates into

$$\left(\frac{h^0_{22}(x)}{h^0_{21}(x)}\right)' > 0 \quad \text{(B.1)}$$

for any $x \in [0,1]$, where the order of functions is without loss of generality. As in Section 2, we assume that the dictator plays a pure strategy conditional on θ_j.

[42] The argument is analogous to the proof of Theorem 1 in Edlin and Shannon (1998).

Lemma 4. *For any* $t_i \in [0,1]$, *the recipient's posterior probability of belief transmission conditional on* t_i *is equal to the ex ante probability of belief transmission.*

Proof. We need to show that for any $t_i \in [0,1]$, $\chi(t_i) = \chi$. Denoting by *NT* the event of no belief transmission, we have

$$\begin{aligned}\chi(t_i) &= 1 - \Pr[NT \mid t_i] = 1 - \frac{(1-\chi)h_1(t_i \mid NT)}{h_1(t_i)} \\ &= 1 - \frac{(1-\chi)\sum_{\kappa=1,2} h_{2,\kappa}^0(t_i)p_\kappa}{h_1(t_i)} = 1 - \frac{(1-\chi)h_3^0(t_i)}{h_1(t_i)} = \chi,\end{aligned} \tag{B.2}$$

where the second equality is by Bayes rule, the third equality is by Assumption A6 and the law of total probability, and the fifth equality is by Assumption A5. ■

Proof of Lemma 1. Given (12) and Lemma 4, the recipient's third-order belief conditional on observing transfer t_i is

$$H_{3,pr}(x \mid t_i) = \chi H_1(x) + (1-\chi)\sum_{\kappa=1,2} H_{2,\kappa}^0 p_{\kappa|t_i}, \tag{B.3}$$

where $p_{\kappa|t_i}$ is the updated probability that the dictator holds SOB $H_{2,\kappa}^0$ conditional on no belief transmission and transfer t_i. Then (since χ does not depend on t_i by Lemma 4), we obtain

$$\frac{\partial H_{3,pr}(x \mid t_i)}{\partial t_i} = (1-\chi)(H_{21}^0(x) - H_{22}^0(x))\frac{\partial p_{1|t_i}}{\partial t_i}. \tag{B.4}$$

By Bayes rule

$$p_{1|t_i} = \frac{h_j(t_i \mid h_{21}^0)p_1}{h_j(t_i \mid h_{21}^0)p_1 + h_j(t_i \mid h_{22}^0)(1-p_1)} = \frac{h_{21}^0(t_i)p_1}{h_{21}^0(t_i)p_1 + h_{22}^0(t_i)(1-p_1)}, \tag{B.5}$$

where the last equality is due to Assumption A6. Taking the derivative of $p_{1|t_i}$ with respect to t_i and substituting it into (B.4) we obtain

$$\begin{aligned}&\frac{\partial H_{3,pr}(x\,|\,t_i)}{\partial t_i}=(1-\chi)(H_{21}^0(x)-H_{22}^0(x))\\&\times\frac{p_1(1-p_1)\left(\dfrac{dh_{21}^0(t_i)}{dt_i}h_{22}^0(t_i)-\dfrac{dh_{22}^0(t_i)}{dt_i}h_{21}^0(t_i)\right)}{(h_{21}^0(t_i)p_1+h_{22}^0(t_i)(1-p_1))^2}.\end{aligned} \tag{B.6}$$

Note that by Assumption A7 this derivative always exists. Consider the right-hand side of (B.6). We have $H_{21}^0(x)-H_{22}^0(x)>0$ for $x\in(0,1)$, since the strict MLRP (B.1) implies strict FOSD of H_{22}^0 over H_{21}^0. Besides, (B.1) yields $\dfrac{dh_{22}^0(t_i)}{dt_i}h_{21}^0(t_i)-\dfrac{dh_{21}^0(t_i)}{dt_i}h_{22}^0(t_i)>0$ for any $t_i\in[0,1]$. Consequently,

$$\frac{\partial H_{3,pr}(x\,|\,t_i)}{\partial t_i}<0 \tag{B.7}$$

for any $x\in(0,1)$ and $t_i\in[0,1]$, which implies the statement of the lemma. ■

Lemma 5. $\dfrac{\partial S_{i,pr}^I(t_i,\theta_j)}{\partial t_i}\le\dfrac{\partial S_{i,pub}^I(t_i,\theta_j)}{\partial t_i}$ *if* $t_i\le\theta_j$ *and* $\beta_i\ge\alpha_i$, *or if* $t_i\ge\theta_j$ *and* $\alpha_i\ge\beta_i$, *with a strict inequality if, in addition,* $0<t_i<1$.

Proof. To avoid notational confusion, let us prove the claim of the proposition for a given value of transfer $t_i=\tilde{t}$.

Let us first consider the intentional surprise in the PUBLIC treatment $S_{i,pub}^I(\tilde{t},\theta_j)$. Given (7) we have

$$\begin{aligned}&S_{i,pub}^I(\tilde{t},\theta_j)=\alpha_i\int_0^{\tilde{t}}(\tilde{t}-x)h_{3,pub}(x,\theta_j)dx-\beta_i\int_{\tilde{t}}^1(x-\tilde{t})h_{3,pub}(x,\theta_j)dx\\&=\alpha_i\int_0^{\tilde{t}}H_{3,pub}(x,\theta_j)dx+\beta_i\int_{\tilde{t}}^1H_{3,pub}(x,\theta_j)dx-\beta_i\left(1-\tilde{t}\right),\end{aligned} \tag{B.8}$$

where the last line is obtained by integration by parts as in (A.3). Taking the derivative we get

$$\frac{\partial S^I_{i,pub}(\tilde{t},\theta_j)}{\partial \tilde{t}} = (\alpha_i - \beta_i) H_{3,pub}(\tilde{t},\theta_j) + \beta_i. \tag{B.9}$$

We have

$$\begin{aligned} &H_{3,pub}(x,\theta_j) = H_1(x,\theta_j) = \chi H_1(x,\theta_j) + (1-\chi) H^0_3(x,\theta_j) \\ &= H_{3,pr}(x,\theta_j) = H_{3,pr}(x \mid t_i = \theta_j), \end{aligned} \tag{B.10}$$

where the first equality is by (10), the second by Assumption A5, and the fourth by Assumption A9.

This, together with (B.9), yields

$$\frac{\partial S^I_{i,pub}(\tilde{t},\theta_j)}{\partial \tilde{t}} = (\alpha_i - \beta_i) H_{3,pr}(\tilde{t} \mid t_i = \theta_j) + \beta_i. \tag{B.11}$$

At the same time, applying integration by parts, we have the following expression for the intentional surprise in the PRIVATE treatment $S^I_{i,pr}(\tilde{t},\theta_j)$ (generally given by (7)):

$$\begin{aligned} &S^I_{i,pr}(\tilde{t},\theta_j) = \alpha_i \int_0^{\tilde{t}} (\tilde{t} - x) h_{3,pr}(x \mid \tilde{t}) dx - \beta_i \int_{\tilde{t}}^1 (x - \tilde{t}) h_{3,pr}(x \mid \tilde{t}) dx \\ &= \alpha_i \int_0^{\tilde{t}} H_{3,pr}(x \mid \tilde{t}) dx + \beta_i \int_{\tilde{t}}^1 H_{3,pr}(x \mid \tilde{t}) dx - \beta_i \left(1 - \tilde{t}\right). \end{aligned} \tag{B.12}$$

Taking the derivative yields

$$\begin{aligned} &\frac{\partial S^I_{i,pr}(\tilde{t},\theta_j)}{\partial \tilde{t}} = (\alpha_i - \beta_i) H_{3,pr}(\tilde{t} \mid \tilde{t}) + \beta_i \\ &+ \alpha_i \int_0^{\tilde{t}} \frac{\partial H_{3,pr}(x \mid \tilde{t})}{\partial \tilde{t}} dx + \beta_i \int_{\tilde{t}}^1 \frac{\partial H_{3,pr}(x \mid \tilde{t})}{\partial \tilde{t}} dx. \end{aligned} \tag{B.13}$$

Subtracting (B.11) from (B.13) we arrive at

$$\frac{\partial S^{I}_{i,pr}(\tilde{t},\theta_j)}{\partial \tilde{t}} - \frac{\partial S^{I}_{i,pub}(\tilde{t},\theta_j)}{\partial \tilde{t}} = D_1 + D_2, \qquad \text{(B.14)}$$

where

$$D_1 = (\alpha_i - \beta_i)(H_{3,pr}(\tilde{t} \mid \tilde{t}) - H_{3,pr}(\tilde{t} \mid \theta_j)), \qquad \text{(B.15)}$$

$$D_2 = \alpha_i \int_0^{\tilde{t}} \frac{\partial H_{3,pr}(x \mid \tilde{t})}{\partial \tilde{t}} dx + \beta_i \int_{\tilde{t}}^{1} \frac{\partial H_{3,pr}(x \mid \tilde{t})}{\partial \tilde{t}} dx. \qquad \text{(B.16)}$$

We have $D_1 \le 0$ by Lemma 1 and initial conditions. At the same time, (B.7) implies that $D_2 \le 0$, with a strict inequality if $0 < \tilde{t} < 1$ (given Assumption A1). Consequently, the LHS of (B.14) is weakly negative, being strictly negative if $0 < \tilde{t} < 1$. ■

Corollary 1. $\frac{\partial U_{i,pr}(t_i,\theta_j)}{\partial t_i} \le \frac{\partial U_{i,pub}(t_i,\theta_j)}{\partial t_i}$ *if* $t_i \le \theta_j$ *and* $\beta_i \ge \alpha_i$, *or if* $t_i \ge \theta_j$ *and* $\alpha_i \ge \beta_i$, *with a strict inequality if, in addition,* $0 < t_i < 1$.

Proof. Given (9) we have

$$U_{i,pr}(t_i,\theta_j) - U_{i,pub}(t_i,\theta_j) = \lambda_2 E_i \left[S^{I}_{i,pr}(t_i,\theta_j) - S^{I}_{i,pub}(t_i,\theta_j) \right]. \qquad \text{(B.17)}$$

This together with Lemma 5 leads to the claim. ■

Corollary 2. *For any* t' *and* $t"$ *so that* $t'' < t' \le \theta_j$ *and* $\beta_i \ge \alpha_i$, *or* $\theta_j \le t'' < t'$ *and* $\alpha_i \ge \beta_i$, *it holds:*

$$U_{i,pub}(t',\theta_j) - U_{i,pub}(t'',\theta_j) > U_{i,pr}(t',\theta_j) - U_{i,pr}(t'',\theta_j).$$

Proof. The claim follows from the fact that $U_i(t',\theta_j) - U_i(t'',\theta_j) = \int_{t''}^{t'} \frac{\partial U_i(x,\theta_j)}{\partial x} dx$ and Corollary 1. ■

Lemma 6. $U_{i,pr}(t_i,\theta_j) > U_{i,pub}(t_i,\theta_j)$ *if* $0 < t_i < \theta_j$, *and* $U_{i,pr}(t_i,\theta_j) \le U_{i,pub}(t_i,\theta_j)$ *if* $t_i \ge \theta_j$.

Proof. To avoid notational confusion, let us prove the claim of the lemma for a given value of transfer $t_i = \tilde{t}$. It follows from (B.8) and (B.10) that

$$S^I_{i,pub}(\tilde{t},\theta_j) = \alpha_i \int_0^{\tilde{t}} H_{3,pr}(x \mid t_i = \theta_j)dx + \beta_i \int_{\tilde{t}}^1 H_{3,pr}(x \mid t_i = \theta_j)dx - \beta_i\left(1-\tilde{t}\right). \tag{B.18}$$

Subtracting (B.18) from (B.12) yields

$$\begin{aligned} &S^I_{i,pr}(\tilde{t},\theta_j) - S^I_{i,pub}(\tilde{t},\theta_j) \\ &= \alpha_i \int_0^{\tilde{t}} \left(H_{3,pr}(x \mid \tilde{t}) - H_{3,pr}(x \mid \theta_j)\right)dx + \beta_i \int_{\tilde{t}}^1 \left(H_{3,pr}(x \mid \tilde{t}) - H_{3,pr}(x \mid \theta_j)\right)dx. \end{aligned} \tag{B.19}$$

It follows from (B.19), Lemma 1 and Assumption A1 that $S^I_{i,pr}(\tilde{t},\theta_j) - S^I_{i,pub}(\tilde{t},\theta_j) > 0$ if $0 < \tilde{t} < \theta_j$, and $S^I_{i,pr}(\tilde{t},\theta_j) - S^I_{i,pub}(\tilde{t},\theta_j) \le 0$ if $\tilde{t} \ge \theta_j$. This, together with (B.17), leads to the claim. ■

Lemma 7. *If* $0 < t^*_{i,pub}(\theta_j) \le \theta_j$ *and* $\beta_i \ge \alpha_i$, *or if* $\theta_j \le t^*_{i,pub}(\theta_j) < 1$ *and* $\alpha_i \ge \beta_i$, *then* $t^*_{i,pr}(\theta_j) < t^*_{i,pub}(\theta_j)$.

Proof. For notational simplicity, let us suppress argument θ_j in the functions of optimal transfers $t^*_{i,pr}(\theta_j)$ and $t^*_{i,pub}(\theta_j)$. Let us first show the weak inequality $t^*_{i,pr} \le t^*_{i,pub}$ under the assumed conditions. Suppose to the contrary that $t^*_{i,pub} < t^*_{i,pr}$. Then, given the initial conditions there can exist only the following cases:

Case 1: $0 < t^*_{i,pub} < t^*_{i,pr} \le \theta_j$ and $\beta_i \ge \alpha_i$, or $\theta_j \le t^*_{i,pub} < t^*_{i,pr}$ and $\alpha_i \ge \beta_i$.

Case 2: $0 < t^*_{i,pub} < \theta_j < t^*_{i,pr}$ and $\beta_i \ge \alpha_i$.

Let us prove that both cases are contradictory.

Case 1:

By Corollary 2 it then follows

$$U_{i,pub}(t^*_{i,pr},\theta_j)-U_{i,pub}(t^*_{i,pub},\theta_j)>U_{i,pr}(t^*_{i.pr},\theta_j)-U_{i,pr}(t^*_{i.pub},\theta_j). \quad \text{(B.20)}$$

At the same time, given that $t^*_{i,pr}$ and $t^*_{i,pub}$ are the optimal choices in the respective treatments, we have

$$U_{i,pr}(t^*_{i,pr},\theta_j)-U_{i,pr}(t^*_{i,pub},\theta_j)\geq 0, \quad \text{(B.21)}$$

$$U_{i,pub}(t^*_{i,pr},\theta_j)-U_{i,pub}(t^*_{i,pub},\theta_j)\leq 0. \quad \text{(B.22)}$$

This contradicts (B.20).

Case 2:

In this case, by Lemma 6

$$U_{i,pr}(t^*_{i,pub},\theta_j)>U_{i,pub}(t^*_{i,pub},\theta_j), \quad \text{(B.23)}$$

$$U_{i,pr}(t^*_{i,pr},\theta_j)\leq U_{i,pub}(t^*_{i,pr},\theta_j). \quad \text{(B.24)}$$

At the same time, since $t^*_{i,pub}$ is the optimal transfer in the PUBLIC treatment, it holds

$$U_{i,pub}(t^*_{i,pr},\theta_j)\leq U_{i,pub}(t^*_{i,pub},\theta_j). \quad \text{(B.25)}$$

It follows from (B.23), (B.24) and (B.25) that

$$U_{i,pr}(t^*_{i,pr},\theta_j)<U_{i,pr}(t^*_{i,pub},\theta_j), \quad \text{(B.26)}$$

contradicting to $t^*_{i,pr}$ being the optimal transfer in the PRIVATE treatment.

Thus, we have come to contradiction in all possible cases when $t^*_{i,pub}<t^*_{i,pr}$, hence

$$t_{i,pr}^* \le t_{i,pub}^*. \tag{B.27}$$

Moreover, this inequality is strict since $0 < t_{i,pub}^* < 1$ by assumption. Indeed, in this case FOC for $t_{i,pub}^*$ is satisfied, i.e.,

$$\frac{\partial U_{i,pub}}{\partial t_i}(t_{i,pub}^*, \theta_j) = 0. \tag{B.28}$$

By Corollary 1 it then follows

$$\frac{\partial U_{i,pr}}{\partial t_i}(t_{i,pub}^*, \theta_j) < 0. \tag{B.29}$$

This means that $U_{i,pr}(t_i, \theta_j)$ is strictly decreasing at $t_i = t_{i,pub}^*$, implying together with (B.27) that $t_{i,pr}^* < t_{i,pub}^*$. ■

Proof of Proposition 2. Given that $0 < t_{i,pub}^*(\theta_j) < 1$ (the existence of interior solutions in the PUBLIC treatment is ensured by Lemma 2), the following first-order condition for maximizing expected utility in the PUBLIC treatment must be satisfied (suppressing the argument in $t_i^*(\theta_j)$ for notational simplicity):

$$\begin{aligned} \frac{\partial U_{i,pub}}{\partial t_i}(t_{i,pub}^*, \theta_j) &= -m_i'(1-t_{i,pub}^*) + \alpha_i \int_0^{t_{i,pub}^*} h_2(x \mid \theta_j)dx + \beta_i \int_{t_{i,pub}^*}^1 h_2(x \mid \theta_j)dx \\ &= (\alpha_i - \beta_i) H_2(t_{i,pub}^* \mid \theta_j) + \beta_i - m_i'(1-t_{i,pub}^*) = 0, \end{aligned} \tag{B.30}$$

where the first equality is by (11) and (8). It follows from the last equality and Assumption A2 that

$$(\alpha_i - \beta_i)\left(H_2(t_{i,pub}^* \mid \theta_j) - \frac{1}{2} \right) \ge 0. \tag{B.31}$$

Consequently, given that θ_j is the median of $H_2(\cdot|\theta_j)$ by (A.8), if $t^*_{i,pub} > \theta_j$, then $\alpha_i \geq \beta_i$, and if $t^*_{i,pub} < \theta_j$, then $\beta_i \geq \alpha_i$. In these cases, $t^*_{i,pr} < t^*_{i,pub}$ by Lemma 7. At the same time, if $t^*_{i,pub} = \theta_j$, then the necessary condition for Lemma 7 is satisfied for any α_i and β_i. Consequently, $t^*_{i,pr} < t^*_{i,pub}$ as well. ■

APPENDIX C. Robustness of Proposition 2 to false consensus

One can show that the analytical results obtained in Section 3.1 are robust to the presence of false consensus, for which we found evidence in Experiment 1. In particular, we show below that if dictators care about the recipients' beliefs and, at the same time, do have a false consensus bias in their SOBs formation, then the same prediction for the difference in transfers between the treatments is obtained as under the consistent formation of beliefs (i.e., as under Assumptions A5 and A6, which we lift here).

The false consensus typically implies that own behavior is considered as representative for the whole population (Ross et al. (1977)). In terms of our model, this means that one's own transfer serves as a signal about others' transfers and, more importantly, about others' representative expectations. This can be expressed in the form of the assumption that, if no information about the recipient's belief is available, the dictator's SOB is formed in the same way as if he got a direct signal θ_j equal to his optimal transfer t^*_i:

$$\tilde{H}_{ij}(x|t^*_i) = H_{ij}(x|\theta_j = t^*_i), \tag{C.1}$$

where $\tilde{H}_{ij}(\cdot|t^*_i)$ is the cdf of the SOB under false consensus. Besides, we assume that the recipient is aware of this fact, i.e., in the PRIVATE treatment her third-order belief is (setting for simplicity $\chi = 0$)

$$H_{jij,pr}(x|t_i) = E_j[\tilde{H}_{ij}(x|t^*_i = t_i)] = E_j[H_{ij}(x|\theta_j = t_i)]. \tag{C.2}$$

This can be justified by the false consensus arising not from some irrational unconscious bias, but rather from a lack of dictator's information, whereby each dictator treats his transfer as the only signal to predict the behavior of others. Such kind of 'rational' false consensus was introduced by Dawes (1989).

Assumption A3 and (C.2) jointly imply that $H_{jij,pr}(x \mid t_i)$ exhibits a strict FOSD in t_i:

$$\frac{\partial H_{jij,pr}(x \mid t_i)}{\partial t_i} < 0 \tag{C.3}$$

for any $x \in (0,1)$ and $t_i \in [0,1]$. Hence, we get the same result as in Lemma 1 for the previous case of rational beliefs. The main difference is that before the positive correlation between transfers and beliefs was driven by the internal consistency of *ex ante* beliefs assumed by recipients (Assumption A6), while in the false consensus case beliefs are shaped by transfers directly.

The FOSD property of the third-order belief allows for the same line of reasoning by comparing the treatments as in the rational case. In the PRIVATE treatment the recipient updates her third-order belief so that it follows the observed transfer, since she believes that the dictator is uninformed and, thus, his SOB is subject to the false consensus effect. In contrast, the (presumed) false consensus effect does not affect beliefs in the PUBLIC treatment, since there we have common knowledge about the actual dictator's SOB (formed by signal θ_j). This leads to a smaller scope of attribution of intentions in the PRIVATE treatment, decreasing the dictator's corresponding motivations and yielding a smaller transfer relatively to the level in the PUBLIC treatment. The formal proofs in this case follow the similar lines as in Appendix B.[43]

We conclude that our analytical predictions from Section 3.1 and our hypothesis from Section 3.2 are robust to the presence of false consensus.

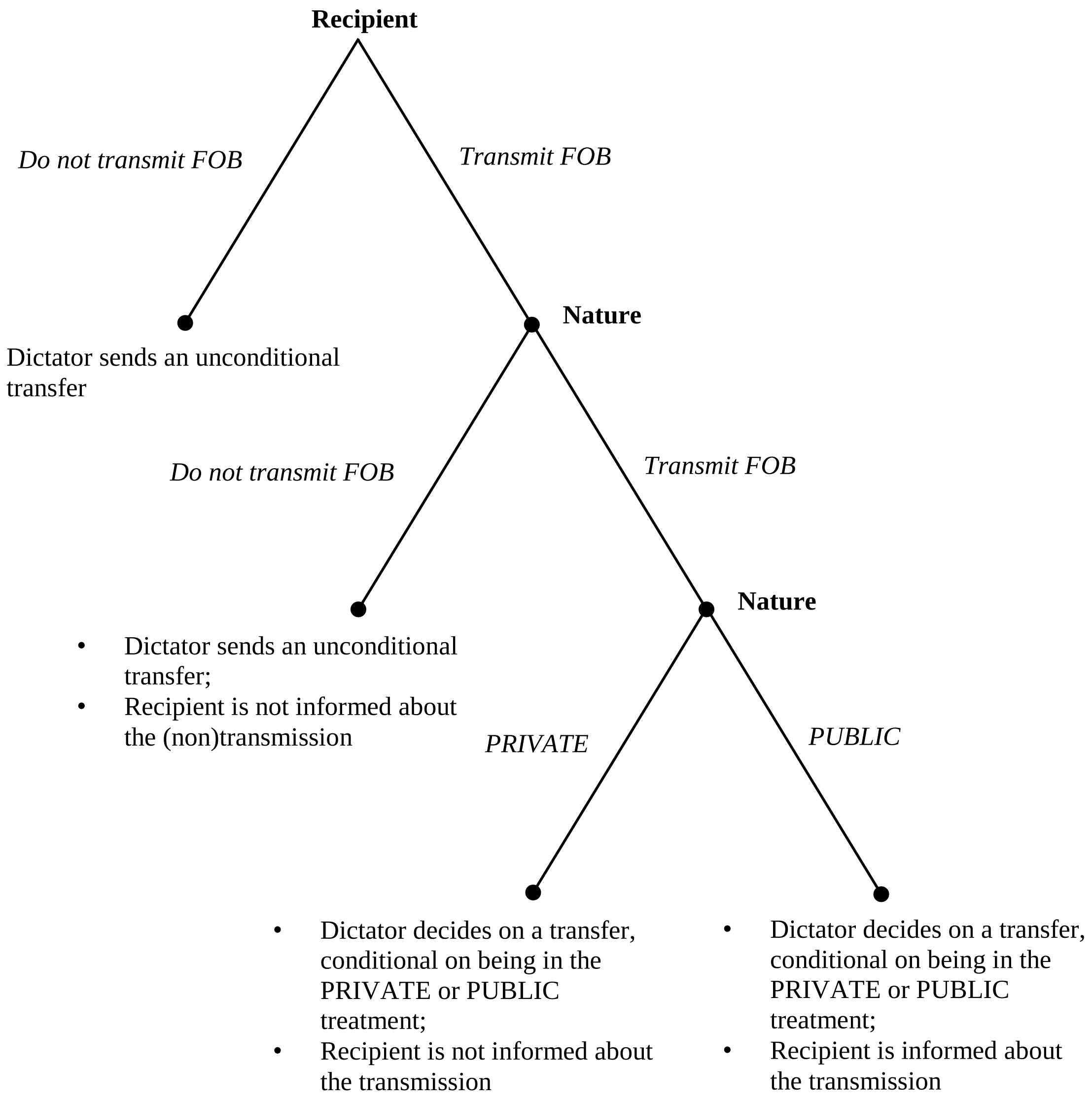
[43] In particular, the result of Lemma 1 is implied by (C.3). The subsequent proof follows the same line of arguments as the proofs in Appendix B starting with Lemma 5 (with the only exception that the result of (B.10) is directly implied by (C.2)).

REFERENCES

R.M. Dawes, Statistical Criteria for Establishing a Truly False Consensus Effect, J. Exp. Soc. Psychol., 25 (1989) 1–17.

L. Ross, D. Greene, P. House, The "False Consensus Effect": An Egoistic Bias in Social Perception and Attribution Processes, J. Exp. Soc. Psychol., 13 (1977) 279–301.

APPENDIX D. Sequence of actions in Part 2 of Experiment 4



APPENDIX E. Experimental instructions

E.1 Experiment 1

Below you find instructions for Experiment 1 translated from German.

Dictators' instructions

General information

Welcome to our experiment. In this experiment you can earn money. You will receive your payoff against the attached receipt, which we ask you to keep.

You are not allowed to speak with other participants during the session. If you have any questions please raise your hand, the experimenter will come to help you. If you violate these rules, we will have to exclude you from the experiment and all payments.

Decision situation

In this experiment all participants are randomly divided into Participants A and Participants B, and each participant is randomly matched with another person. **You are Participant A, the other person is Participant B.**

Each pair receives an endowment of 14 Euro. Then you have to decide about how this sum should be divided between you and Participant B. This means, you determine your own amount and the amount of Participant B so that

Your payoff=14 Euro – amount of Participant B

Payoff of Participant B= amount of Participant B

Prior to your decision your matched Participant B will be asked to guess how much of the 14 Euro, on average, Participant A will send to Participant B. You will be informed about the guess of your matched Participant B first after your decision about the division of the sum is made. However, you can set the amount of Participant B to

depend on the possible guesses of Participant B. The payoff-relevant amount of Participant B is then the amount that you chose for the actual guess of Participant B.
Participant B does not know that you will be informed about his guess and that you can condition your decision on it.

Participant B can get an additional payoff by his guess. The Participant B, whose guess is the closest to the actual average amount received by Participants B, wins an additional bonus of 8.00 Euro. If several participants are closest, then the person who gets the bonus is determined randomly.
Take your time and make sure you understand these instructions.
All decisions and payoffs are confidential. No other participant will get to know your payoffs.
Moreover, no participant will get to know during or after the experiment which other participant he or she was assigned to.

Form for Participant A

Please indicate your decision.

If the (rounded) guess of my Participant B about his amount is the following (see inputs in this column)…	… then I give him the following amount:
0,00 €	_ _ , _ _ €
0,50 €	_ _ , _ _ €
1,00 €	_ _ , _ _ €
1,50 €	_ _ , _ _ €
2,00 €	_ _ , _ _ €
2,50 €	_ _ , _ _ €
3,00 €	_ _ , _ _ €
3,50 €	_ _ , _ _ €
4,00 €	_ _ , _ _ €
4,50 €	_ _ , _ _ €
5,00 €	_ _ , _ _ €
5,50 €	_ _ , _ _ €
6,00 €	_ _ , _ _ €
6,50 €	_ _ , _ _ €
7,00 €	_ _ , _ _ €
7,50 €	_ _ , _ _ €
8,00 €	_ _ , _ _ €
8,50 €	_ _ , _ _ €
9,00 € and more	_ _ , _ _ €

Post-experimental questionnaire

Finally, we would like to ask you a few questions.

Age: ____ years

Gender: (female/male)

Field of study: ______

Semester: ______

Mother tongue: ______

Do you know the decision situation from a previous experiment? (Yes/No)

What do you think is the amount that Participant A should send to Participant B? [0.00-14.00 Euro] _ _ , _ _ Euro.

What do you think is the average amount that Participant A sends to Participant B? [0.00-14.00 Euro] _ _ , _ _ Euro.

What do you think is the average guess of all Participants B about the amount sent by Participant A to Participant B? [0.00-14.00 Euro] _ _ , _ _ Euro.

Recipients' instructions

General information

Welcome to our experiment. In this experiment you can earn money. You will receive your payoff against the attached receipt, which we ask you to keep.

You are not allowed to speak with other participants during the session. If you have any questions please raise your hand, the experimenter will come to help you. If you violate these rules, we will have to exclude you from the experiment and all payments.

Decision situation

In this experiment all participants are randomly divided into Participants A and Participants B, and each participant is randomly matched with another person. **You are Participant B, the other person is Participant A.**

Each pair receives an endowment of 14 Euro. Then Participant A has to decide about how this sum should be divided between himself and Participant B. This means, he determines his own amount and your amount so that

Payoff of Participant A=14 Euro – your amount
Your payoff = your amount

Before Participant A makes the decision, you will be asked to guess how much of the 14 Euro, on average, Participant A will send to Participant B.
You can get an additional payoff by your guess. The Participant B, whose guess is the closest to the actual average amount received by Participants B, wins an additional bonus of 8.00 Euro. If several participants are closest, then the person who gets the bonus is determined randomly.
Take your time and make sure you understand these instructions.
All decisions and payoffs are confidential. No other participant will get to know your payoffs.
Moreover, no participant will get to know during or after the experiment which other participant he or she was assigned to.

Form for Participant B

What do you believe is the average amount that Participants B will get?
Please state a value from [0.00 - 14.00 Euro]:

___ ___ , ___ ___ Euro.

Post-experimental questionnaire

Finally, we would like to ask you a few questions.
Age: _____ years
Gender: (female/male)
Field of study: _______

Semester: ______

Mother tongue: ______

Do you know the decision situation from a previous experiment? (Yes/No)

What do you think is the amount that Participant A should send to Participant B? [0.00-14.00 Euro] _ _ , _ _ Euro.

What is the amount that you would have sent in the role of Participant A? [0.00-14.00 Euro] _ _ , _ _ Euro.

What do you think? What does your matched Participant A think is the amount that Participants B expect on average? [0.00-14.00 Euro] _ _ , _ _ Euro.

E.2 Experiment 2

Below you find instructions for Experiment 2 translated from German. The instructions for treatments PUBLIC and PRIVATE differ only in the sentences marked in the text.

Dictators' instructions

General information

Welcome to the experiment! In this experiment you can earn money. How much you can earn depends on your decisions. You will receive an amount of 2.50 Euro for your participation that will be paid out irrespective of the decisions in the experiment.

From now on please do not communicate with other participants. If you have a question, please raise your hand! We will come to your desk and answer your question. If you violate these rules, we will have to exclude you from the experiment and all payments.

Decision situation

In this experiment, two participants are randomly matched. One participant is randomly assigned the role of Participant A, the other is randomly assigned the role of Participant B.

You are Participant A, the other person is Participant B.

You receive an endowment of 10 Euro. From this endowment, you can send any amount to Participant B. Payoffs are calculated as follows:

Your payoff = 10 Euro – amount sent

Payoff of B = Amount sent

Prior to your decision, Participant B will be asked to guess how much of the 10 Euro, on average, Participant A will send to Participant B. You will be informed about Participant B's guess before you decide on the amount to be sent.

[Treatment PUBLIC] After Participant B made the guess he/she will be informed that you know his/her guess before you choose the amount to be sent.

[Treatment PRIVATE] Participant B will not be informed that you know his/her guess.

Participant B can achieve an additional payoff through his/her guess. The Participant B whose guess is closest to the actual average will win an amount of 8 Euro.

Take your time and make sure that you understand these instructions. All decisions and payoffs are confidential. No participant will get to know your payoffs, and you will receive the money in a closed envelope when you leave the laboratory.

Moreover, no participant will get to know during or after the experiment which other participant he or she was assigned to.

Post-experimental questionnaire

Finally, we would like to ask you a few questions.

Age: ____ years

Gender: (female/male)

Faculty: (business/economics, law, medicine, arts and humanities, mathematics and natural sciences, human sciences, no student)

Semester: ______

Mother tongue: ______

Do you know the decision situation from a previous experiment? (Yes/No)

What do you think is the amount that Participant A should send to Participant B? [0.00-10.00 Euro] _ _ , _ _ Euro.
What do you think is the average amount that Participant A sends to Participant B? [0.00-10.00 Euro] _ _ , _ _ Euro.
What do you think is the average guess of all Participants B about the amount sent by Participant A? [0.00-10.00 Euro] _ _ , _ _ Euro.

Recipients' instructions

Instructions for Participant B were identical in the PUBLIC and the PRIVATE treatments. In the PUBLIC treatment, Participants B additionally received the following information displayed on the computer screens after they had typed in their guesses: "Participant A will be informed about your guess of ____Euro before he decides on the amount sent to you."

General information

Welcome to the experiment! In this experiment you can earn money. How much you can earn depends on your decisions. You will receive an amount of 2.50 Euro for your participation that will be paid out irrespective of the decisions in the experiment.
From now on please do not communicate with other participants. If you have a question, please raise your hand! We will come to your desk and answer your question. If you violate these rules, we will have to exclude you from the experiment and all payments.

Decision situation

In this experiment, two participants are randomly matched. One participant is randomly assigned the role of Participant A, the other is randomly assigned the role of Participant B.
You are Participant B, the other person is Participant A.
Participant A receives an endowment of 10 Euro. From this endowment, he or she can send any amount to you. Payoffs are calculated as follows:

Payoff of A = 10 Euro – amount sent

Your Payoff = Amount sent

Prior to Participant A's decision, you will be asked to guess how much of the 10 Euro, on average, Participant A will send to Participant B.

You can achieve an additional payoff through your guess. The Participant B whose guess is closest to the actual average will win an amount of 8 Euro.

Take your time and make sure that you understand these instructions. All decisions and payoffs are confidential. No participant will get to know your payoffs, and you will receive the money in a closed envelope when you leave the laboratory.

Moreover, no participant will get to know during or after the experiment which other participant he or she was assigned to.

Post-experimental questionnaire

Finally, we would like to ask you a few questions.

Age: ____ years

Gender: (female/male)

Faculty: (business/economics, law, medicine, arts and humanities, mathematics and natural sciences, human sciences, no student)

Semester: ______

Mother tongue: ______

Do you know the decision situation from a previous experiment? (Yes/No)

What do you think is the amount that Participant A should send to Participant B? [0.00-10.00 Euro] _ _ , _ _ Euro.

What is the amount that you would have sent in the role of Participant A? [0.00-10.00 Euro] _ _ , _ _ Euro.

What do you think? What does your matched Participant A think is the amount that Participants B expect on average? [0.00-10.00 Euro] _ _ , _ _ Euro.

E.3 Experiment 3

Below you find instructions for Experiment 3 translated from German. Dictators and recipients obtained the same instructions.

General Information

Welcome to the experiment! In this experiment you can earn money. How much you earn depends on your decisions and the decisions of other participants.

Your payoff and your decisions are confidential. No participant will know, during or after the experiment, whom he has interacted with and how much other participants earned. Your decisions are thus anonymous.

The experiment consists of two parts. First, you receive the instructions for the first part of the experiment. After the first part is finished, you receive the instructions for the second part of the experiment.

After the experiment is finished, one of the two parts of the experiment will be randomly chosen for all participants. The payoffs resulting from the decisions of the participants in this randomly chosen experimental part will then be paid out.

From now on please do not communicate with other participants. If you have a question concerning the experiment please raise your hand! We then come to you and answer your question. If you violate these rules, we will have to exclude you from the experiment and all payments.

Instructions for the first part

In this experiment all participants are randomly divided into Participants A and Participants B, and each participant is randomly matched with another person so that each Participant A is matched with a Participant B.

All participants get an amount of 2.50 Euro, which is paid to them independently of their decisions in the first part of the experiment.

Participant A receives an endowment of 14 Euro from us. He can then send any amount from this endowment to Participant B. The payoffs are as follows:

Payoff of Participant A = 14 Euro – amount sent

Payoff of Participant B = amount sent

Participant B will be informed about the amount, which was sent to him by Participant A, only at the end of the experiment, i.e., after its second part.

Before Participant A takes the decision about the amount to send, Participant B is asked to guess the amount that Participants A are going to send on average to Participants B. Also Participant A, prior to his decision about the amount to send, is asked about what he thinks is the average guess submitted by Participants B.

Participants A as well as Participants B can earn additional payoffs through their guesses, because the best guess among Participants A as well as among Participants B will be rewarded with 8 Euro. This reward will be paid at the end of the experiment in any case, independently of which experimental part is randomly chosen.

This is the end of the instructions for the first experimental part. Take your time and make sure that you understand these instructions. If you still have questions, please raise your hand and one of the experimenters will come to you.

Instructions for the second part

In the second part of the experiment all participants keep the roles (Participant A and Participant B), which were assigned to them before the first part.

Each Participant A is matched with a new Participant B. It is ensured that no Participant A is matched with the same Participant B, with whom he has been already matched in the first part.

Participant A receives an amount of 2.50 Euro, which is paid to him independently of the decisions in the second part of the experiment.

Participant B can decide in the beginning of the second part whether his guess from the first part (his guess about the average amount sent to Participants B by Participants A) may be transmitted to his currently matched Participant A.

If Participant B decides that her guess may be transmitted, she gets an amount of 2.50 Euro, which is paid to her independently of the decisions in the second part of the experiment.

If Participant B decides against the transmission of his guess, he then does not get the amount of 2.50 Euro.

As in the first part of the experiment, Participant A receives from us an endowment of 14 Euro. He can then send any amount from this endowment to Participant B. The decision of Participant A about the amount to be sent to Participant B proceeds as follows:

First, Participant A is informed whether his matched Participant B has allowed to transmit his guess to him. There are two possible cases:

1) If Participant B has not allowed to transmit his guess, then Participant A submits only one amount to send. The payoffs then are as follows:

   Payoff of Participant A = 14 Euro – amount sent

   Payoff of Participant B = amount sent

2) If Participant B has allowed to transmit his guess, then Participant A has to submit "conditional" amounts to send.

   "Conditional" amounts to send are amounts which are set depending on possible guesses of Participant B. This means that Participant A submits an amount to send for each possible (rounded) guess of Participant B (see the table below).

If the (rounded) guess of my Participant B about the average amount sent is the following (see inputs in this column)…	… then I send my Participant B the following amount:
0,00 €	_ _ , _ _ €
0,50 €	_ _ , _ _ €
1,00 €	_ _ , _ _ €
1,50 €	_ _ , _ _ €
2,00 €	_ _ , _ _ €
2,50 €	_ _ , _ _ €
3,00 €	_ _ , _ _ €
3,50 €	_ _ , _ _ €
4,00 €	_ _ , _ _ €
4,50 €	_ _ , _ _ €
5,00 €	_ _ , _ _ €
5,50 €	_ _ , _ _ €
6,00 €	_ _ , _ _ €
6,50 €	_ _ , _ _ €
7,00 €	_ _ , _ _ €
7,50 €	_ _ , _ _ €
8,00 €	_ _ , _ _ €
8,50 €	_ _ , _ _ €
9,00 € and more	_ _ , _ _ €

At this point, Participant A does not yet know the actual guess of Participant B. He is informed about the guess after he has decided about all "conditional" amounts. The amount, which is then relevant for payment, is the amount that Participant A has chosen for the true guess of Participant B. Then, the payoffs are as follows:

Payoff of Participant A = 14 Euro – amount sent for the true guess of Participant B

Payoff of Participant B = amount sent for the true guess of Participant B

This is the end of the instructions for the second experimental part. Take your time and make sure that you understand these instructions. If you still have questions, please raise your hand and one of the experimenters will come to you.

Post-experimental questionnaire

Final questionnaire: Please answer the following questions:

Have you been Participant A or Participant B during the experiment?

Age: ____ years

Gender: (female/male)

In which faculty do you study: (business/economics, law, medicine, arts and humanities, mathematics and natural sciences, human sciences, no student)

Semester: ______

Mother tongue: ______

In how many experiments have you participated before (approximately)?

Do you know the decision situation from a previous experiment? (Yes/No)

E.4 Experiment 4

Below you find instructions for Experiment 4 translated from German. Dictators and recipients obtained the same instructions.

General Information

Welcome to the experiment! In this experiment you can earn money. How much you earn depends on your decisions and the decisions of other participants.

Your payoff and your decisions are confidential. No participant will know, during or after the experiment, whom he has interacted with and how much other participants earned. Your decisions are thus anonymous.

The experiment consists of two parts. First, you receive the instructions for the first part of the experiment. After the first part is finished, you receive the instructions for the second part of the experiment. All participants in this experiment receive identical instructions.

After the experiment is finished, one of the two parts of the experiment will be randomly chosen for all participants. The payoffs resulting from the decisions of the participants in this randomly chosen experimental part will then be paid out to the participants.

From now on please do not communicate with other participants. If you have a question concerning the experiment please raise your hand! We then come to you and answer your question. If you violate these rules, we will have to exclude you from the experiment and all payments.

Instructions for the first part

In this experiment all participants are randomly divided into Participants A and Participants B, and each participant is randomly matched with another person so that each Participant A is matched with a Participant B.

All participants get an amount of 2.50 Euro, which is paid to them independently of their decisions in the first part of the experiment.

Participant A receives an endowment of 14 Euro from us. He can then send any amount from this endowment to Participant B. The payoffs are as follows:

Payoff of Participant A = 14 Euro – amount sent

Payoff of Participant B = amount sent

Participant B will be informed about the amount, which was sent to him by Participant A, only at the end of the experiment, i.e., after its second part.

Before Participant A makes the decision about the amount to send, Participant B is asked to guess the amount that Participants A are going to send on average to Participants B. Also Participant A, prior to his decision about the amount to send, is asked about what he thinks is the average guess submitted by Participants B.

Participants A as well as Participants B can earn additional payoffs through their guesses, because the best guess among Participants A as well as among Participants B will be rewarded with 8 Euro. This reward will be paid at the end of the experiment in any case, independently of which experimental part is randomly chosen.

This is the end of the instructions for the first experimental part. Take your time and make sure that you understand these instructions. If you still have questions, please raise your hand and one of the experimenters will come to you.

Instructions for the second part

In the second part of the experiment all participants keep the roles (Participant A and Participant B), which were assigned to them before the first part.

The second part of the experiment consists of 2 rounds. At the beginning of each round in this part, each Participant A is matched with a new Participant B. It is ensured that no Participant A is matched with the same Participant B, with whom he has been already matched in the first part or in the previous round.

Participant A receives an amount of 2.50 Euro, which is paid to him independently of the decisions in the second part of the experiment.

Each round of the second part consists of two phases:

Phase 1: Transmission of the guess

Participant B can decide whether **his guess from the first part** (his guess about the average amount sent to Participants B by Participants A**) may be transmitted to the Participants A that he will be matched with in the second part**.

If Participant B decides that his guess may be transmitted, she also gets an amount of 2.50 Euro, which is paid to him independently of the decisions in the second part of the experiment.

If Participant B decides against the transmission of his guess, he then does not get the amount of 2.50 Euro.

Phase 2: Decision about the amount sent and the information of the participants

In both rounds of the second part of the experiment, Participant A receives from us an endowment of 14 Euro. He can then send any amount from this endowment to the Participant B he is matched with in the current round.

If Participant B in Phase 1 has allowed to transmit his guess, **it is nevertheless uncertain if the Participant A he is matched with in the current round will be actually informed about the guess**. In this case, it is randomly determined whether or not Participant A will be informed about the guess.

With 50% probability, the guess will be transmitted; in this case Participant A knows the guess of the Participant B he is matched with in the current round.

Otherwise, the guess will not be transmitted; in this case Participant A does not know the guess of the Participant B he is matched with in the current round.

If the guess was transmitted to Participant A, **Participant B will be informed with 50% probability at the end of the experiment if Participant A in fact knew her**

guess in the particular round. In all other cases, Participant B will not be informed about this.

If Participant B has allowed to transmit his guess and the guess was actually transmitted, Participant A submits conditional amounts to send in the current round:

- The amount (0 to 14 Euro) he wants to send in case Participant B is informed at the end of the experiment whether Participant A knew his guess.
- The amount (0 to 14 Euro) he wants to send in case Participant B is not informed at the end of the experiment whether Participant A knew his guess.

The payoffs are as follows, conditional on the actual information of Participant B:

Payoff of Participant A = 14 Euro – conditional amount sent

Payoff of Participant B = conditional amount sent

If Participant B has not allowed to transmit her guess or he has allowed but the guess has not been transmitted, Participant A submits the amount to send (0 to 14 Euro) to the Participant B in the current round, without knowing his guess.

The payoffs are then as follows:

Payoff of Participant A = 14 Euro – amount sent

Payoff of Participant B = amount sent

At the end of the experiment, Participant B will be informed about the amounts sent in both rounds.

If the second part is relevant for the payoffs of the participants, one of the two rounds of this part is randomly chosen. The payoffs resulting from the decisions of the participants in this round will then be paid out to the participants.

This is the end of the instructions for the second experimental part. Take your time and make sure that you understand these instructions. If you still have questions, please raise your hand and one of the experimenters will come to you.

Post-experimental questionnaire

Final questionnaire: Please answer the following questions:

Have you been Participant A or Participant B during the experiment?

Age: ____ years

Gender: (female/male)

In which faculty do you study: (business/economics, law, medicine, arts and humanities, mathematics and natural sciences, human sciences, no student)

Semester: ______

Mother tongue: ______

In how many experiments have you participated before (approximately)?

Do you know the decision situation from a previous experiment? (Yes/No)

APPENDIX F. Additional results

Experiment 2

Table F.1 lists the results of Tobit models with the amount sent by the dictator as the dependent variable. In Model 1, we only include a dummy variable for the PUBLIC treatment. Its coefficient is positive and highly significant, corroborating our treatment effect. In Model 2, we additionally include the expectation of the matched recipient which turns out to be insignificant while the PUBLIC dummy is largely unaffected. Finally, the strong treatment effect and the lack of an impact of the matched recipient's belief remain unchanged when we additionally include variables capturing the demographic backgrounds of the subjects (see Model 3).[44]

Table F.1

Determinants of dictator transfers in Experiment 2.

No.	1	2	3
Dependent Variable	Transfer	Transfer	Transfer
Model	Tobit	Tobit	Tobit
PUBLIC	0.982**	0.998**	1.017**
	[0.400]	[0.399]	[0.395]
Recipient's expectation		0.121	0.138
		[0.144]	[0.141]
Age			-0.009
			[0.056]
Female			-0.113
			[0.401]
Business student			-0.152
			[0.414]
Dictator game known			-0.799**
			[0.399]
Constant	0.353	0.076	-0.047
	[0.303]	[0.452]	[1.633]
Observations	127	127	127
Log-Likelihood	-218.4	-218.0	-215.8

Tobit models are calculated to account for the share of observations with zero transfers. Standard errors are given in brackets. ** denotes significance at the 5%-level. 'Age' is the participant's age in years. The dummy variable 'Female' takes the value of 1 if the participant is female. 'Business student' is a dummy variable equal to 1 if the participant is enrolled at the faculty of economics and business. Finally, 'Dictator game known' is a dummy variable indicating whether the subject knew the decision situation in advance.

[44] The only significant impact of the demographical background is found for subjects who previously knew the decision situation of the dictator game – the negative and significant coefficient indicates that these subjects decrease their transfers compared to subjects without previous experience. Nobody participated twice in our experiments.

Experiment 3

Table F.2 contrasts the results of Experiment 1 with the results from the relevant Part 2 of Experiment 3. In Experiment 3, dictators transferred on average 2.24 Euro conditioned on the true FOB of their matched recipient. This accounts for 16% of the endowment which is less than the 24% from Experiment 1, perhaps due to the fact that Experiment 1 was conducted in the classroom while Experiment 3 was ran in the laboratory which might create larger social distance between participants.[45]

Table F.2
Comparison of results: Experiment 1 versus Experiment 3

	Experiment 3	**Experiment 1**
Average transfer	2.24 Euro (16% of endowment)	3.25 Euro (23% of endowment)
Between-subject correlation coefficient of transfers with guesses	0.054, *p*=0.619	-0.017, *p*=0.821
Share of dictators who vary conditional transfers	71.6%	77.5%
Share of dictators with a statistically significant (at the 5% level) within-subject correlation	58.0%	53.9%
Share of dictators with a significantly positive correlation	43.2%	37.7%
Share of dictators with a significantly negative correlation	14.8%	16.2%
Share of transfers strictly above guesses	23.9%	27.2%
Average size of positive regression coefficients	0.36	0.58
Average size of negative regression coefficients	0.29	0.53
Share of non-monotonic subjects with a significant within-subject correlation	15.7%	30.1%

Experiment 4

Table F.3 lists Tobit models to account for the share of observations with zero transfers, while including random effects for the dictators. The dependent variable in these models is the amount sent in Part 2 of Experiment 4 for the cases in which the

[45] The average amount sent in Part 1 of Experiment 3 accounts for 2.84 Euro or 20.3% of the endowment.

dictator in fact observed the FOB of his matched recipient.[46] The coefficient of the dummy variable for the PUBLIC treatment is significant in all specifications, while the coefficient of the recipient's expectation is insignificant in Models 2 and 3. With respect to the effect of the demographic variables, and unlike in the model for Experiment 2 reported in Table F.1, 'Age' has a significant negative coefficient, indicating that older dictators transfer less. Contrary to Model 3 in Table F.1, the dummy 'Dictator game known' has an (insignificant) positive coefficient. All other demographic variables are (as before) insignificant.

Table F.3

Determinants of dictator transfers in Experiment 4 (Part 2, case of belief transmission).

No.	1	2	3
Dependent Variable	Transfer	Transfer	Transfer
Model	Tobit	Tobit	Tobit
PUBLIC	0.529***	0.529***	0.531***
	[0.182]	[0.182]	[0.183]
Recipient's expectation		0.183	0.268
		[0.173]	[0.168]
Age			-0.203**
			[0.089]
Female			0.608
			[0.642]
Business student			0.069
			[0.652]
Dictator game known			0.981
			[0.648]
Constant	2.257***	1.625**	5.343**
	[0.360]	[0.702]	[2.253]
Observations	126	126	124
Log-Likelihood	-229.0	-228.4	-221.3

Tobit models with random effects on the level of dictators are calculated to account for the share of observations with zero transfers. Standard errors are given in brackets. *** and ** denote significance at the 1%- and 5%-level, respectively. 'Age' is the participant's age in years. The dummy variable 'Female' takes the value of 1 if the participant is female. 'Business student' is a dummy variable equal to 1 if the participant is enrolled at the faculty of economics and business. Finally, 'Dictator game known' is a dummy variable indicating whether the subject knew the decision situation in advance.

[46] The average amount sent in Part 1 of Experiment 4 is 3.03 Euro or 21.6% of the endowment.