

Zhurakhovska, Lilia

Working Paper

Strategic trustworthiness via unstrategic third-party reward: An experiment

Preprints of the Max Planck Institute for Research on Collective Goods, No. 2014/6

Provided in Cooperation with:

Max Planck Institute for Research on Collective Goods

Suggested Citation: Zhurakhovska, Lilia (2014) : Strategic trustworthiness via unstrategic third-party reward: An experiment, Preprints of the Max Planck Institute for Research on Collective Goods, No. 2014/6, Max Planck Institute for Research on Collective Goods, Bonn

This Version is available at:

<https://hdl.handle.net/10419/102671>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



**Strategic Trustworthiness via
Unstrategic Third-party Reward
– An Experiment**

Lilia Zhurakhovska





Strategic Trustworthiness via Unstrategic Third-party Reward – An Experiment

Lilia Zhurakhovska

June 2014

STRATEGIC TRUSTWORTHINESS VIA UNSTRATEGIC THIRD-PARTY REWARD – AN EXPERIMENT

LILIA ZHURAKHOVSKA

UNIVERSITY OF COLOGNE,
UNIVERSITY OF ERLANGEN-NUREMBERG,
MAX PLANCK INSTITUTE FOR RESEARCH ON COLLECTIVE GOODS, BONN

The paper investigates the introduction of an institution, in form of an impartial authority (third party), into a two-person situation. The impartial authority can reward a stranger for acting according to a desired behavioral norm. The reward is costly for the authority and her behavior cannot be strategic, i.e., it cannot lead to higher earnings for her in a later stage. A trust game followed by a helping game is implemented. The trustee in the trust game becomes the recipient in the helping game. This paper demonstrates that positive strong indirect reciprocity can exist in one-shot settings (helpers transfer more money to their recipients the more these are trustworthy) and that positive strong indirect reciprocity is not diminishing if the reward can be anticipated. Furthermore, the positive strong indirect reciprocity is correctly anticipated and leads to higher trustworthiness in the treatment.

Key words: strong indirect reciprocity, third-party reward, trust game, helping game, anticipation, norms

JEL: C90, D03, D63

I. INTRODUCTION

"Virtually every commercial transaction has within itself an element of trust, certainly any transaction conducted over a period of time. It can be plausibly argued that much of the economic backwardness in the world can be explained by the lack of mutual confidence."

(Kenneth J. Arrow, 1972)

The concept of trust has always played a crucial role in economics (cf. Arrow 1972). It can be a money-maximizing strategy to trust someone more the higher that person's incentives are to be trustworthy. In repeated settings, one is trustworthy (or cooperates) because one expects others to reciprocate (positively or negatively) one's behavior, and thus mutual trustworthiness leads to higher earnings for the actors. Strategically motivated reciprocity can be carried out by the person who is directly affected by someone's act himself (on direct reciprocity, see, e.g., Fehr, Gächter 2000; Andreoni and Miller 1993) or by another person (on indirect reciprocity, see, e.g., Rockenbach and Milinski 2006; Seinen and Schram 2006; Engelmann and Fischbacher 2009). In modern societies, however, many contacts are anonymous, indirect, and rarely (if ever) repeated. One example of such transactions is anonymous online trades between private persons via platforms such as www.ebayclassifieds.com/ or www.craigslist.org/. Consequently, it is important to study how trustworthiness can be enhanced by the anticipation of a reward or punishment by an impartial stranger who does not have an incentive to reciprocate (strong indirect reciprocity). I.e., in this paper, "positive strong indirect reciprocity" is defined as a third-party reward in situations in which any strategic concerns for the third parties can be excluded.¹ Consider the following example, which illustrates the meaning and meaningfulness of anticipated (positive) "strong indirect reciprocity" for trustworthiness. Think of a politician who runs for a post. The voters invest their trust in the candidate by voting for him. Assume an institutional design in which reelection is not possible, and thus in which the politician has no strategic incentives to reciprocate that trust. Typically, not everything that is done behind the curtains of a public institution is transparent to the public. Therefore, it is not easy fully to judge how well the politician does his job. No image, after the politician has been elected, his institution implements a new policy. This new policy implies that reports on the effort provided by the politicians are publicly available. Imagine you learn from the reports that the politician does a good job. Imagine further that you are the head of a bank in which the very same politician applies for a loan. First of all, beyond all money-maximizing concerns from your perspective: would you give him the loan with a higher probability, compared to a situation in which

¹ Carpenter and Matthews (2004) and Carpenter et al. (2004) relate to "strong indirect reciprocity" as "social reciprocity". Camerer and Fehr (2004, p. 56) simply define "Reciprocity means that people are willing to reward friendly actions and to punish hostile actions although the reward or punishment causes a net reduction in the material payoff of those who reward or punish."

you had not learned that he has reciprocated his voters' trust (positive strong indirect reciprocity)? Second, would the probability of your giving him the loan be even higher if you knew that while the politician was doing his job he did not know that the report might become publically available (motivational crowding out/in of strong indirect reciprocity)? Third, would the politician put more effort in his work if he were able to anticipate that this could be rewarded by others who have no extrinsic incentives to do so (higher trustworthiness in anticipation of positive strong indirect reciprocity)? Finally, would more voters elect the politician if they knew in advance that he would have an incentive to invest in his good reputation in anticipation of a reward by others (higher trust)? This paper attempts to answers all these questions. It provides evidence that positive strong indirect reciprocity exists; it is anticipated by potential recipients; and it can change these recipients' previous behavior in an efficient way.

To the best of my knowledge, this is the first paper explicitly to study the effect of anticipated positive strong indirect reciprocity on trustworthy behavior.² Following the call by Almenberg et al. (2011),³ it is one of the few papers to study positive strong indirect reciprocity. Additionally, in this paper, the impartial stranger has a richer strategy set to condition her⁴ rewards on the history of her beneficiary than in previous studies. So far, only one study analyzes whether there is a crowding out of strong indirect positive reciprocity, if a potential recipient of this reward can act strategically (Stanca et al. 2009). Interestingly, in contrast to Stanca et al. (2009), the present study finds no crowding out.

In the present experimental paper, a trust game is implemented and it is followed by a helping game⁵ with a different group composition, i.e., the trustee of the trust game in one group becomes a receiver in a helping game in a different group. The helper has the possibility to reward a co-player conditional on his performance in the trust game. In the *Baseline*, subjects first play the trust game and receive the instructions for the helping game only afterwards, while in the *Anticipation* treatment, subjects are informed about the content of both games at the beginning of the experiment. Since the experiment is one-shot, any strategic concerns for the helpers are ruled out in both treatments.

Nonetheless, many helpers make positive transfers to trustees and send significantly more money as the trustees' return transfers increase. Helpers apparently care more about socially desirable behavior than about the motives behind trustees' transfers, i.e., on average, helpers' transfers are the same, regardless of whether the helping game is announced or not. Trustees anticipate helpers' behavior if the helping game

² In Stanca (2009) and Stanca et al. (2009), the anticipation effect of a reward by an impartial stranger on the behavior of a first-mover (instead of an effect on a second-mover in anticipation of a reward by an impartial third-party) is reported. However, this aspect is not the focus of these papers.

³ Almenberg et al. (2011) state "*While costly punishment has received the lion's share of attention, costly rewarding also plays an important role in human prosociality.*" They continue: "*[A] sizeable amount of evidence exists for the importance of rewarding in human cooperation. Yet the reward-based analog to third party punishment, where I reward you in an anonymous one-shot interaction because you have cooperated with somebody else, remains largely unexplored.*" (Almenberg et al. 2011, p. 75, p. 77).

⁴ Throughout the paper, the female form "she" is used for the third parties (and for investors, i.e., for players A in the experimental design) and the male form "he" for the other players.

⁵ The helping game is, in fact, a simple dictator game with an efficiency factor. The name of the game, which is often used in the literature, might be misleading, since the game does not necessarily have anything to do with help for a needy person.

is announced, i.e., the absolute level of return transfers as well as the relative level to the investments are higher in the *Anticipation* treatment as compared to the *Baseline*. Investments, on average, do not differ between the *Anticipation* treatment and the *Baseline*.

The remainder of the paper is organized as follows: first, an overview of the relevant literature is presented; afterwards, the design of the experiment is explained. Next, hypotheses are explored. The results and the statistical analysis of the data are presented in the penultimate section. In the last section, the conclusions are drawn.

II. LITERATURE

I am aware of only three studies on positive strong indirect reciprocity.⁶ In contrast to the present study, none of these three studies analyzes strategic versus unstrategic trustworthiness (due to an anticipation of a reward by an impartial stranger). The most closely related papers to the present study are Stanca (2009) and Stanca et al. (2009).⁷ In both studies, the return transfer in a one-shot variant of the trust game either comes from the recipient of the investment or from a stranger.⁸ In contrast to the present study, both papers focus on the strong indirect reciprocity (return transfers) and do not analyze in detail the effect on behavior of a player in *anticipation* of strong indirect reciprocity (a change in investment). Stanca (2009), however, does report not finding significant differences in the investments. This means that in his experiment the strategic motives for the players awaiting direct or indirect reciprocity do not seem to matter. Stanca et al. (2009) hypothesize that the motives behind the reciprocated action can crowd out strong indirect reciprocity (rewards by impartial strangers). The treatment difference here is whether the first mover knows that the second stage will follow. The results confirm their hypothesis, i.e., if the strong indirect reciprocator knows that the first mover was aware of the second stage, she transfers a smaller amount compared to a situation in which the second game was not announced. Notably, the results in the present study are not in line with their hypothesis. Almenberg et al. (2011) implement a one-shot dictator game where a player can transfer either none, half, or all of her endowment to another player, and a third party can either punish or reward the dictator. Furthermore, the number of recipients, the achievable share, and the effectiveness of the reward and punishment given by the third party are varied across treatments. In all treatments, dictators are aware of the presence of the third party. Thus, an effect of anticipated strong indirect reciprocity cannot be studied

⁶ In fact, the basic design in Almenberg et al. (2011) is very similar to Stanca (2009) and Stanca et al. (2009). In all three studies presented in this section, player A transfers an amount of money to player B which can be observed and rewarded and/or sanctioned by player C.

⁷ Here only the papers on the topic of downstream/social indirect reciprocity (A acts towards B and C acts as a reaction to this in a certain way towards B) are discussed, since these papers are most relevant for the present study. Nevertheless, it is important to mention that there are also interesting papers on generalized/upstream indirect reciprocity (A acts towards B and B acts as a reaction to that in a certain way towards A). Notable examples are Dufwenberg et al. (2001); Boyd and Richardson 1989; Greiner and Levati (2005); Güth et al. (2001).

⁸ He calls it a gift-exchange game.

in their setting. The authors' main findings show that selfish behavior is punished while generous behavior is rewarded, and that rewards are at least as common as punishments.

Quite a few studies analyze negative strong indirect reciprocity. Only one recent paper examines the anticipation effect of punishment. Balafoutas, Nikiforakis, and Grechenig (2014) demonstrate, in a one-shot, three-player taking game, that taking rates decrease in anticipation of unstrategic punishment by an impartial third party. Furthermore, their paper shows that third-party punishment significantly decreases if counter-punishment directed towards the third party is allowed. However, the focus of their paper lies on counter-punishment, which could be interpreted as an emotional reaction or revenge. Carpenter and Matthews (2004) run a repeated public-goods experiment with punishment. In one treatment, only members of the own group can be punished, while in the other treatment, members of the own as well as of another group can be punished. The authors find evidence for the existence of negative strong indirect reciprocity, i.e., members of stranger-groups are punished. Fehr and Fischbacher (2004) suggest that Carpenter and Matthews (2004) "*could not rule out third-party punishment for reasons of self-interest*". There is also a strong disciplining component in their design, i.e., punishing someone should lead to more norm compliance in future periods. Fehr and Fischbacher (2004) find strong evidence for third-party punishment in their one-shot, three-person dictator experiment. Bernhard et al. (2006) run one-shot dictator games with third-party punishment in Papua New Guinea. They find that in-group members are avenged more than out-group members, while the affiliation of the punished person does not play a role for the punishers' decisions. On the contrary, norm violators expect to be punished less if the third party belongs to their peer group than if she belongs to a different group.

III. DESIGN

This section is divided into three subsections. In the first subsection, the experimental design is explained; in the second, the motivation behind this design is discussed; and finally, the experimental procedures followed in the experiment are reported.

III.1. GAME AND TREATMENTS

The game consists of two parts and subjects were aware of that.⁹ In the *Baseline*, they receive specific information about the content of each part only immediately before playing the relevant part of the experiment. In the *Anticipation* treatment, instructions for both parts are handed out at the beginning of the

⁹ In total, there are three parts. Part 1 and Part 2 are described in this section. Part 3 is a standard risk aversion elicitation experiment (Holt and Laury 2002). The results of the risk aversion measure are not reported, as these are not relevant for the present study.

experiment.¹⁰ Subjects are explicitly told that they cannot lose the money they have earned in a previous part in any of the subsequent parts. In the experiment, the experimental currency unit (ECU) is used. All instructions are read out aloud by the experimenter to achieve common knowledge about the procedure.¹¹

At the beginning of the experiment, each subject is randomly assigned one of the three roles A, B, or C. Players keep their roles for the two parts of the experiment. The roles A and B are assigned to 11 subjects each and the role C is assigned to 2 players per session. The distribution of the roles is not made explicit to the subjects. They only know which roles there are, their own role, and, at the relevant point in time, the role of their co-player. The group composition differs between the parts, i.e., players from part 1 do not meet in part 2. The game is played only once.

Part 1:

In part 1, the reduced trust game (TG) (first introduced by Berg et al. 1995) is played by two players, A and B, who move sequentially. The players are endowed with $E^{TG} = 100$ ECU each. At first, player A (from now on called investor) decides how many ECU she wants to send to B (called trustee from now on). Her transfer $t_A^{TG} = X$ can be 0, 10, 20, 30, 40, 50, or 60 (from now on called player A's investment).

The investment is tripled by the experimenter and then transferred to B. In case of a positive investment, B can make a return transfer to A. His return transfer is $t_B^{TG} = Y * X$, where Y can be 0, 1, 2, or 3. The decision by player B is elicited via the strategy method (Selten 1967), i.e., B decides about Y for each possible X.¹² Table 1 displays player B's strategy table. After player B has made his decision, all players are informed about their payoffs from the trust game (π^{TG_i}). These are:

$$\pi_A^{TG} = E^{TG} - t_A^{TG} + t_B^{TG} \text{ for player A and } \pi_B^{TG} = E^{TG} + t_A^{TG} * 3 - t_B^{TG} \text{ for player B.}$$

Players C do not take any decision in the TG, nor are they informed about the decisions of the other players. Players C receive a fixed payoff of $\pi_C^{TG} = 100$.¹³

¹⁰ A treatment that comprises only the trust game was run as well. The results of the *trust game-only* treatment do not differ significantly from the results of the trust game in the *Baseline*; therefore, they are not reported here.

¹¹ See section VIII.1 in the Appendix for an English translation of the instructions.

¹² Brandts and Charness (2011) show that, if the focus lies on the comparison of decisions within strategies, using the strategy method can be problematic. For the comparison between treatments, the main limitation of that method is that the "strategy method provides a lower bound for testing for treatment effects" (p. 392). A similar argument is made in Fischbacher et al. (2012).

¹³ Please note that, in case of zero investment, $\pi_A^{TG} = \pi_B^{TG} = \pi_C^{TG} = 100$.

TABLE 1

Experimental Design – Trustees' Strategy

| In case player A has sent me... | | | |
|--|---|-------|-----------------|
| ... 0 Taler , my income has thereby increased by 0 Taler . | | | |
| ... 10 Taler and my income has thereby increased by 30 Taler , I will now send | <input type="radio"/> nothing | (0) | to player A. |
| | <input type="radio"/> the transfer | (10) | |
| | <input type="radio"/> double the transfer | (20) | |
| | <input type="radio"/> triple the transfer | (30) | |
| ... 20 Taler and my income has thereby increased by 60 Taler , I will now send | <input type="radio"/> nothing | (0) | to player A. |
| | <input type="radio"/> the transfer | (20) | |
| | <input type="radio"/> double the transfer | (40) | |
| | <input type="radio"/> triple the transfer | (60) | |
| ... 30 Taler and my income has thereby increased by 90 Taler , I will now send | <input type="radio"/> nothing | (0) | to player A. |
| | <input type="radio"/> the transfer | (30) | |
| | <input type="radio"/> double the transfer | (60) | |
| | <input type="radio"/> triple the transfer | (90) | |
| ... 40 Taler and my income has thereby increased by 120 Taler , I will now send | <input type="radio"/> nothing | (0) | to player A. |
| | <input type="radio"/> the transfer | (40) | |
| | <input type="radio"/> double the transfer | (80) | |
| | <input type="radio"/> triple the transfer | (120) | |
| ... 50 Taler and my income has thereby increased by 150 Taler , I will now send | <input type="radio"/> nothing | (0) | to player A. |
| | <input type="radio"/> the transfer | (50) | |
| | <input type="radio"/> double the transfer | (100) | |
| | <input type="radio"/> triple the transfer | (150) | |
| ... 60 Taler and my income has thereby increased by 180 Taler , I will now send | <input type="radio"/> nothing | (0) | to player A. |
| | <input type="radio"/> the transfer | (60) | |
| | <input type="radio"/> double the transfer | (120) | |
| | <input type="radio"/> triple the transfer | (180) | |

In the box, the screen for the elicitation of trustees' choices via the strategy method is depicted. In the first column, the trustee can see how high the investment could have been. In the second column, the participant sees a radio button, on which he can click – for each possible investment $X > 0$, he can choose how much he wants to send back to his investor, i.e., he can choose for each investment $X > 0$ his return transfer $X \cdot Y$, where Y can be 0 ("nothing"), 1 ("the transfer"), 2 ("double the transfer"), or 3 ("triple the transfer"). Only the transfer decision for the relevant situation will become payoff-relevant.

Part 2:

For the helping game (HG) (similar to the dictator game, as in Forsythe et al. 1994), played in part 2, new groups with two players are formed. Each group consists of a player A and a co-player. The randomly

selected co-player is either player C or player B, who has not been matched with this player A in the trust game, i.e., absolute stranger matching is implemented. Player A (from now on called the helper) is endowed with $E^{HG}_A = 100$ ECU. She can transfer any natural number of ECU (t^{HG}_A) from 0 to 100 to her co-player. The transfer is tripled by the experimenter and then transferred to the relevant co-player. Player A's decision is elicited using the strategy method (Selten 1967).¹⁴ She has to make a decision for every possible composition, i.e., she has to state how much she wants to transfer in case her co-player is player B or in case her co-player is player C. Additionally, supposing B is her co-player, she can make her decision conditional on the history of the player B in the TG, i.e., she can make her transfer conditional on any possible return transfer t^{TG}_B of her co-player to his investor given any possible prior investment t^{TG}_A her co-player might have received. This means that the helper makes 26 transfer decisions. Table 2 gives an overview of helpers' strategy space. Only the payoff-relevant transfer decision is realized, based on whether the co-player is in fact a player B or a player C and, if applicable, based on player B's actual return transfer t^{TG}_B and on the investment he has received. After the decisions have been made, the players are informed about the actual group composition and the relevant transfer.¹⁵ The payoffs in the helping game (π^{HG}_i) are:

$$\pi^{HG}_A = E^{HG}_A - t^{HG}_A \text{ for the helper and } \pi^{HG}_{co-player} = t^{HG}_A * 3 \text{ for the co-player.}$$

¹⁴ Again, one could argue that the strategy method prompts subjects to take different decisions for different situations. However, the results show that helpers indeed condition their transfers on the relative return transfers of the trustees, but less on the investments that trustees receive. Furthermore, helpers' transfers are more strongly correlated with trustees' history in the *Anticipation* treatment compared to the *Baseline*, which again cannot be explained by the use of the strategy method.

¹⁵ Subjects are informed in the instructions that the players who are not randomly selected to become the helper's co-players, can earn money m in an additional task. It is not made explicit in the instructions how many subjects have to perform that task, nor how much money can be earned in that task. Only the players who are chosen to perform that task receive additional instructions for the task on their computer screen. The task is to count the number of zeros in tables that consisted of 150 randomly ordered zeros and ones (similar to the task used in Abeler et al. 2011). Each correct answer increases m by 50 ECU. The payoff for these players is therefore $\pi^{HG}_{other player} = m$. In each session, 10 of the helpers were matched with one player B each and one helper was matched with one player C. Therefore, one B player and one C player were not chosen to become a co-player of the helper.

TABLE 2

Experimental Design – Helpers' Strategy

| |
|--|
| <p>1st Screen</p> <div style="border: 1px solid black; padding: 10px; margin: 10px 0;"> <p>If, in Stage 2, my co-player is ...</p> <ul style="list-style-type: none"> • ... a Player C, ... <div style="text-align: right;">... I will now send ____ Taler.</div> • ... a Player B who has been sent 0 Taler in the first experiment, ... <div style="text-align: right;">... I will now send ____ Taler.</div> </div> |
| <p>2nd-6th Screens</p> <div style="border: 1px solid black; padding: 10px; margin: 10px 0;"> <p>If, in Stage 2, my co-player is ...</p> <ul style="list-style-type: none"> • ... a Player B who has been sent X Taler in the first part of the experiment and... <div style="text-align: right;">... sent back 0*X Taler, I will now send ____ Taler.</div> <div style="text-align: right;">... sent back 1*X Taler, I will now send ____ Taler.</div> <div style="text-align: right;">... sent back 2*X Taler, I will now send ____ Taler.</div> <div style="text-align: right;">... sent back 3*X Taler, I will now send ____ Taler.</div> </div> |

In the first box, the first screen of the elicitation of helpers' choices via the strategy method is depicted. Here, the helper can choose how much she wants to send to her co-player in case her co-player is a player C or in case her co-player is a player B who has not received an investment. The second box shows the remaining screens of the helpers' strategy method. Here, the helper can choose how much she wants to transfer to her co-player in case he is a player B and has received an investment of X (X=10 on the 2nd screen, X=20 on the 3rd screen, etc.) and in case he has then sent back Y*X (Y can be 0, 1, 2, or 3). On the actual screens of the players, instead of "X", "0*X", or "1*X" etc., the absolute numbers of the respective transfers are written. The helper can insert in each line a number between 0 and 100. Only the transfer decision for the relevant situation will become payoff-relevant.

After the subjects have finished the experiment, the subjects complete questionnaires regarding their attitudes towards trust, risk, and reciprocity, as well as demographics.

III.2. MOTIVATION FOR THE DESIGN

A restricted version of the trust game is used for reasons of simplification, i.e., this allows helpers to have a complete overview of every possible situation they might be facing. This method allows for testing for strong indirect reciprocity of helpers, i.e., it provides detailed data for helpers' transfers conditional on previous behavior of their co-players.

The inclusion of a player C, who does not take any actions, helps to identify an individual benchmark for the helpers' general willingness to help. Helpers' transfers to C players cannot be interpreted

as a reward for any previous action. These players have the same income as players B, who are passive, i.e., who do not receive an investment and who therefore cannot make a return transfer.

In the experiment, a player A becomes the helper in the helping game. Charness and Rabin (2002) have developed a theory that analyzes disinterested social preferences (non-self-interested distributional preferences). One potential critique is that in the present study the helping game involves a mixture of self-interested and disinterested preferences. Giving each player feedback about the payoffs after each game and controlling for these payoffs in a regression analysis (instead of trying to elicit beliefs about their earnings) helps to disentangle these motives. Furthermore, participation in the first part of the experiment facilitates the understanding of the strategy method in the helping game.

III.3. EXPERIMENTAL PROCEDURE

Experiments were run at the University of Bonn in May 2010 and in January 2012. The experiment was programmed and conducted using z-Tree (Fischbacher 2007). Four sessions with 96 subjects were held, leading to 22 independent observations per treatment for the roles A and B, respectively. Subjects were invited from the University of Bonn using ORSEE software (Greiner 2004) and had no experience with related experiments, i.e., they had neither participated in trust games nor in dictator or helping games in the past. Most of the subjects were students. 19 of the subjects were economics students, 20 were law students. The remaining subjects came from various different disciplines. 44 participants (46%) were female. The average earning was 12.52 Euro.¹⁶ The sessions lasted 70 minutes on average.

IV. HYPOTHESES

In the helping game, under the assumption of pure payoff maximization, the theoretical prediction is zero transfers. A self-interested, payoff-maximizing trustee has no monetary incentive for a positive return transfer, irrelevant of the information about the helping game. A rational, self-interested, payoff-maximizing investor anticipates this and does not invest any points in any treatment. From the theoretical point of view, under the assumption of pure payoff maximization and common knowledge, the unique Nash Equilibrium predicts zero transfers in all games and in all treatments.

However, it has been shown that, in helping games, positive transfers are observed (cf., e.g., Forsythe et al. 1994), and in trust games positive transfers are made in both directions (see, e.g., Berg et al. 1995). The first is explained by social preferences such as warm glow (Andreoni 1990). The latter is

¹⁶ That corresponded approximately to \$15.25 during the first wave of the experiment and approximately to \$15.96 during the second wave of the experiment.

explained by strong direct reciprocity modeled by, e.g., Rabin (1993), Dufwenberg and Kirchsteiger (2004), Falk and Fischbacher (2006), or Levine (1998). These models assume that an actor has a reciprocity parameter and that the person has a positive utility from punishing unkind action and rewarding kind action. Which actions are perceived as kind depends on the particular model. In fact, the intuition of the model by Levine (1998) is not restricted to two-player direct interactions. In the following, this model will be used to derive (most of the) behavioral hypotheses in the present study. In Levine (1998) a player $i = 1, \dots, n$ receives a *direct utility* of u_i and has a coefficient of altruism $-1 < a_i < 1$. Therefore, he receives an *adjusted utility* of

$$v_i = u_i + \sum_{j \neq i} \frac{a_i + \lambda a_j}{1 + \lambda} u_j,$$

where $0 \leq \lambda \leq 1$. In other words, when $\lambda > 0$, a person derives a positive utility from his own direct utility (which can be his payoff) and, in addition, a positive utility from rewarding another person for his altruism, i.e., i 's utility from j 's utility is greater the greater j 's coefficient of altruism is.¹⁷ In the model, a player i maximizes her utility given her preferences and her beliefs about the preferences of her co-players. In the present experiment, the helper does not need to form her own beliefs about the preferences of her co-player, since she can infer them from the trustee's return transfers to his investor. Assuming that at least some helpers have a coefficient of altruism $a_i > 0$ and $\lambda > 0$, it follows:

Hypothesis 1a: The higher a helper's co-player's return transfers in the trust game are, the more help will be transferred.

The second key research question is whether helpers' evaluation of trustees' altruism differs between the *Baseline* and the *Anticipation* treatment. Levine's model does not make specific assumptions about intentions behind the displayed altruism. Therefore, from Levine (1998) it does not follow that helpers' transfers should differ between the *Baseline* and the *Anticipation* treatment. I.e., if helpers in the experiment care mainly about socially desirable behavior, they should not make different transfers in the *Baseline* compared to the *Anticipation* treatment. However, Falk et al. (2008) show that intentions matter.¹⁸ If a trustee anticipates that he will be rewarded by a helper in the *Anticipation* treatment, and therefore increases his return transfer for strategic reasons, his act might be perceived as less kind. Falk et al. (2008) argue that players acting out of intrinsic motives will be rewarded more than other players. This finding suggests that, in the present experiment, helpers' transfers should be higher in the *Baseline* than in the *Anticipation* treatment because, in the *Baseline*, trustees' return transfers cannot be strategically motivated.

¹⁷ Depending on the parameters, the model can also make predictions on negative reciprocity and on unconditional altruism. However, since punishment is not possible in the present experiment, these predictions will not be discussed in this paper. Transfers to players C can be explained by unconditional altruism of helpers.

¹⁸ Some theories on reciprocity incorporate intentions as well (e.g., Rabin 1993, Charness and Rabin 2002).

Similarly, Stanca et al. (2009) find that more strong indirect reciprocity is displayed when strategic motivations can be ruled out. This leads to:

Hypothesis 1b: For every possible return transfer of trustees, helpers' transfers are lower in the *Anticipation* treatment than in the *Baseline*.

Following the model in Levine (1998), one can derive the following prediction about trustees' return transfers in the *Baseline*:

Hypothesis 2a: The trustees in the trust game make higher positive return transfers the higher the investment they receive.

In the *Anticipation* treatment, "*predictive power of the theory is about what we would expect from a signaling model*" (Levine 1998, p. 605). Since the trustee not only cares about being altruistic (reciprocal) to others (to the investor), but also cares about his personal utility (his payoffs), he has to form beliefs about whether signaling that he is altruistic (trustworthy) can be beneficial to him. In the *Baseline*, the expected probability of being rewarded for revealed altruism (trustworthiness) is zero.¹⁹ In the *Anticipation* treatment, however, the trustee knows that a helper has the opportunity to reward him. If at least some trustees expect some helpers to have a coefficient of altruism $a_i > 0$ and $\lambda > 0$, they should expect to receive higher transfers in the helping game the more trustworthy they are, i.e., they anticipate the prediction in Hypothesis 1a. If a positive number of trustees expects their helper's a_i to be large enough to offset the monetary loss from higher return transfers, these trustees have an incentive to make higher return transfers in the *Anticipation* treatment than in the *Baseline*. Consequently, one can derive the following hypothesis:

Hypothesis 2b: The trustees in the trust game make higher positive return transfers in the *Anticipation* treatment than in the *Baseline*.

Costa-Gomes et al. (2012) find a positive correlation between an investors' investment and his optimism about a return transfer. Hence, if at least some investors anticipate the predictions about return transfers, which follows from Levine (1998), this leads to:

Hypothesis 3: Investors will make higher investments in the *Anticipation* treatment than in the *Baseline*.

¹⁹ One could argue that by not announcing the helping game in the *Baseline*, but by informing the players that a second game will follow, some subjects might expect their actions in the trust game to have an influence of their later payoffs. However, the expected probably for that is centennially greater in the *Anticipation* treatment than in the *Baseline*.

V. RESULTS

This section is organized as follows: first, the hypotheses concerning the transfers are tested non-parametrically²⁰ and then robustness of the results with parametric tests including further control variables is checked. The results are presented in the same order as the respective hypotheses.

V.1. HELPERS' TRANSFERS

This section first investigates whether helpers' transfers depend on relative return transfers made by trustees in the trust game and on the investment received by the trustees. Then, treatment differences on helpers' transfers are tested. In addition, it is analyzed which level of the return transfers is particularly strongly rewarded, i.e., what level of return transfers is regarded as especially altruistic and therefore worth to be rewarded.

One quarter of helpers (11 out of 44) always make zero transfers in the helping game (7 in *Baseline* and 4 in *Anticipation*).²¹ As one can see in Figure 1, higher relative return transfers are rewarded by higher helpers' transfers (Spearman's Rho: $r_s=0.4815$, $p=0.0000$; *Baseline*: $r_s=0.4352$, $p=0.0000$; *Anticipation*: $r_s=0.5286$, $p=0.0000$).²² Figure 1 furthermore suggests that helpers' transfers for a particular return transfer do not depend on investments (Spearman's Rho: $r_s=0.0152$, $p=0.6211$; *Baseline*: $r_s=0.0118$, $p=0.7875$; *Anticipation*: $r_s=0.0221$, $p=0.6123$).²³ Overall, transfers are very similar in the *Baseline* and in the *Anticipation* treatment. The simple non-parametric analysis of all average transfers of the helpers yields no significant treatment difference (Mann-Whitney rank-sum $|z|=1.112$, $p\text{-value}=0.2661$).²⁴

²⁰ Throughout this paper, reported p-values are always two-sided, unless stated otherwise.

²¹ The likelihood for a helper to make any positive transfer is not different between the treatments: 1-sided Fisher's exact = 0.244.

²² This finding is further supported by the Wilcoxon signed-rank tests (see Table 5 in the Appendix).

²³ This finding is further supported by the Wilcoxon signed-rank tests (see Table 7 in the Appendix).

²⁴ See Table 8 in the Appendix for all pairwise comparisons.

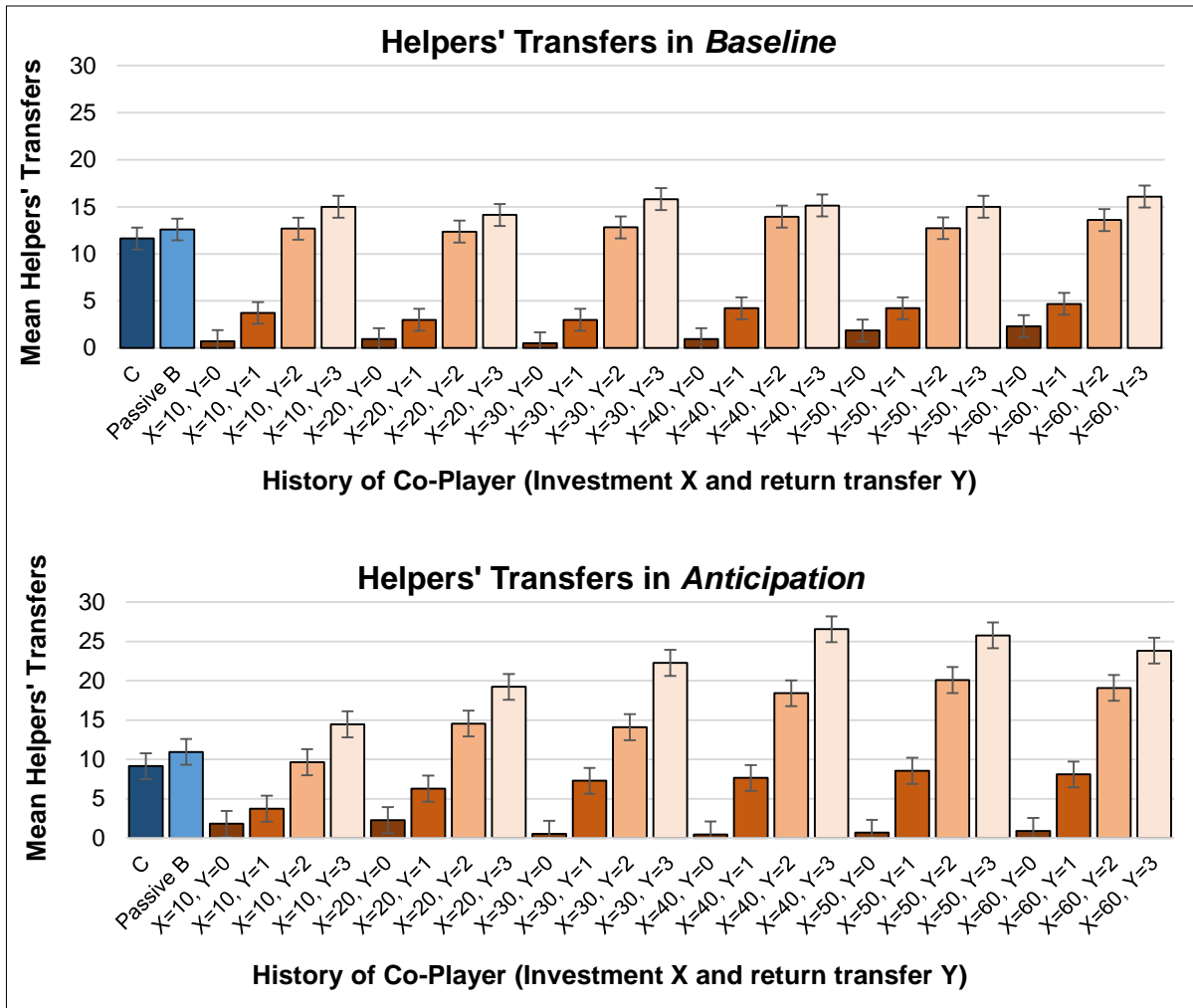


Figure 1

Helpers' Transfers by Treatment

The upper figure displays helpers' transfers in the *Baseline* and the lower figure in the *Anticipation* treatment. On the x-axis, the exact condition is displayed, i.e., one can see if the co-player is either Player C or a passive Player B (no investment and thus no opportunity for a return transfer) or an active Player B who has received an investment of X and has made a relative return transfer of Y (Y=0: trustee keeps full transfer; Y=1: trustee returns transfer and keeps rent; Y=2: equal split; Y=3 full return). On the y-axis, mean helpers' transfers are displayed for the particular situation. Standard errors are indicated.

Table 3
Explaining helpers' transfers – comparison Baseline and Anticipation treatment

| Random effects Tobit regressions ("helpers" as group) | | | | |
|---|---------------------|----------------------|----------------------|----------------------|
| Dependent variable: Helpers' transfers in the helping game to active trustees | | | | |
| | Model 1 | Model 2 | Model 3 | Model 4 |
| <i>Anticipation</i> | 10.05 (10.59) | 5.66 (10.38) | 9.87 (8.52) | -1.24 (9.43) |
| Relative Return Transfers | 16.43*** (.72) | 16.43*** (.72) | 16.41*** (.72) | 14.82*** (1.00) |
| Investment | .12*** (.04) | .12*** (.04) | .12*** (.04) | .044 (.06) |
| <i>Anticipation</i> *Relative Return Transfer | | | | 2.94** (1.37) |
| <i>Anticipation</i> *Investment | | | | .14* (.08) |
| Own Investment | | .03 (.26) | -.23 (.23) | -.23 (.26) |
| Own Profit in TG | | .36** (.17) | .12 (.15) | .12 (.15) |
| Transfer to passive B | | | .34 (.47) | .71 (.45) |
| Transfer to C | | | .71 (.45) | .32 (.47) |
| Constant | -47.22*** (7.99) | -82.30*** (20.41) | -62.71*** (17.12) | -56.65*** (17.25) |
| N | 1056 | 1056 | 1056 | 1056 |
| N of group | 44 | 44 | 44 | 44 |
| P model | <.001 | <.001 | <.001 | <.001 |
| Wald Chi2 | 528.48 | 530.67 | 539.19 | 550.51 |

Random effects Tobit regressions. Standard errors are presented in parentheses. The *Anticipation* dummy equals 1 for all observations of the *Anticipation* treatment, *relative return transfers* controls for the relative return of a trustee (Y) for a given investment, *investment* controls for the investment the trustee has received, *own investment* is the investment the helper has transfers in the trust game himself to his trustee, *Anticipation*relative return transfer* and *Anticipation*investment* are interaction terms, *own profit in TG* controls for the helper's profit from part 1 of the experiment. *Transfer to passive B* and *Transfer to C* are the levels transferred to passive players. Significance at the 10%, 5%, and 1% level is denoted by *, **, and ***, respectively. Left-censored = 577; right-censored = 14.

Using the random effects Tobit regression model, one can confirm the visual impression.²⁵ The dependent variable is the helpers' transfers in the helping game. In no model in Table 1 is the dummy

²⁵ The Tobit regression is used because, in the helping game, helpers' transfers are exogenously restricted with an upper and a lower bound; the lower bound is usually zero-giving. Bardsley (2008) shows that subjects also take money if they have the opportunity in similar situations. In the present setting, this seems plausible, since, as stated before, relative return transfers are rewarded a lot by helpers, while very low relative return transfers lead to very low helpers' transfers and often to transfers of zero. Tobit regressions account for the possibility that (some) subjects might have even taken money instead of giving nothing by controlling for censoring. Moreover, as there are 24 transfer decisions per individual (due to the strategy method – only taking into account transfers to active

variable for the *Anticipation* treatment significant. Model 1 in Table 3 shows that the main determining factor for helpers' transfers is the relative return transfer of trustees, i.e., higher relative return transfers yield higher helpers' transfers.²⁶ The investment the trustees received (resulting in a higher or lower absolute return transfer) has a significant, but very small, positive influence on helpers' transfers. The coefficient is not significant once one controls for an interaction between the treatment and the effect of the investment (see Models 4).

Model 2 controls for helpers' own experience in the trust game. It shows that pure willingness to send positive transfers does not lead to higher helpers' transfers (variable *own investment*). However, there could be a small wealth effect: the more a helper has earned in the previous trust game, the more willing she is to help in the helping game.²⁷ The significance vanishes once one controls for helpers' transfers to passive players (see Model 3), which can be used as a measure of a helpers' individual benchmark of altruism.

In Model 4, an interaction term between the *Anticipation* treatment and the *investment* to the trustee and an interaction term between the *Anticipation* treatment and the *relative return transfer* of the trustee are added. The coefficient of the investment is not significant in this model. On the contrary, coefficients of both interaction terms are significant and positive. The most important result derived from Model 3 in the Tobit regressions is that trustees are treated somewhat differently in the *Anticipation* treatment than in the *Baseline*. In contrast to the prediction in Hypothesis 1b, they receive, on average, for a given history (a particular investment followed by a particular return transfer) a higher transfer from a helper, if they can invest in their good reputation knowing that the helping game will follow. This result contradicts the findings in Stanca et al. (2009). It also calls for a new model of (positive strong indirect) reciprocity in which other factors besides intentions are considered. In the present experiment, helpers seem to care more about socially desirable behavior of trustees than about intrinsically motivated intentions behind the return transfers.²⁸

In the regressions, there is strong left-censoring, which indicates that helpers would possibly like not just to give less money to, but even to take money from greedy trustees. The results from the random effects Tobit regression mainly support Hypothesis 1a, while there is no clear support of Hypothesis 1b. This leads to:

players), random effects models which take individual specific effects into account are in order. The coefficient of the treatment dummy is directly interpretable in the sense that it gives exactly the value of the average marginal effect of the independent variable.

²⁶ The Spearman correlation analysis is used to see if any personal traits of participants influenced their helpers' transfers. Negative reciprocity is negatively correlated with helpers' transfers (Spearman's $\rho = -0.3247$; $p\text{-value} = 0.0315$). Besides that, neither the gender nor any other personality trait is significantly correlated with helpers' transfers.

²⁷ One could also interpret that coefficient as a proxy for generalized/upstream indirect reciprocity – the more a helper has received in return from his own trustee, the more does she help in the helping game. With the help of the regression, one can disentangle that motive from the social/downstream indirect reciprocity motive of the helper, which is captured in the variable *relative return transfers*.

²⁸ One should be cautious in interpreting this result, since in the present experiment helpers arise from the population of investors. Therefore, they might feel they belong to the group of investors. As a result, they might care more about the total earnings resulting from high return transfers of investors than about potential strategic motives of trustees.

Result 1a: The higher the relative return transfers a trustee makes in the trust game, the more he receives from a helper. Generally, the investment that a trustee has received does not seem to determine a helper's transfer to the trustee.

Result 1b: Helpers' transfers are, on average, not lower if the helping game is announced. The transfers are more positively correlated with relative as well as with absolute return transfers, if the helping game is announced.

The strategy of helpers includes passive players (Player C and Player B, who do not receive an investment and can therefore not make any return transfers). Helpers' transfers to passive players can be regarded as a benchmark of how much a helper is willing to transfer to a player who does not have a history.²⁹ Furthermore, it can be identified which distributive norms achieved in the trust game helpers reward (higher transfer than to passive players) and which they punish (lower transfer than to passive players). A relative return transfer of $Y=3$ leaves the trustee with his endowment only and reciprocates the investment completely; a relative return transfer of $Y=2$ leads to an equal split between the trustee and the investor; by a relative return transfer of $Y=1$, the investor is compensated for his investment and earns as much as his endowment would have been without an investment, while the trustee keeps the complete rent from the investment; the least generous possible relative return transfer in the experiment is $Y=0$, i.e., this relative return transfer makes the investor worse off than if he had not made an investment and leaves the trustee with the highest possible earning. A Wilcoxon signed-rank test shows that the passive Players C are not treated differently than passive Players B ($|z|=0.653$, $p = 0.5138$; *Baseline* only: $|z|=0.899$, $p = 0.3688$; *Anticipation* only $|z|=0.196$, $p = 0.8444$). Figure 1 indicates that, for any investment, if a trustee makes a relative return transfer of $Y = 2$, he receives on average the same transfer by the helper as a passive player; for each investment for a relative return transfer of $Y < 2$ he receives a lower transfer than a passive player; and, respectively, for each investment a relative return transfer of $Y=3 > 2$ earns him a higher transfer by a helper as to a passive player. Indeed, this visual impression is supported by Wilcoxon signed-rank tests (see Table 6 in the Appendix) in both treatments.³⁰ Given these observations one can state:

²⁹ Please note that each Player C and each Player B receives an endowment of 100 ECU. Only after receiving an investment from Player A can a Player B become active and thereby increase the investor's and his own payoffs. In case Player B returns the entire transfer, he ends up with the same payoff as a Player C or a passive Player B (with 100 ECU).

³⁰ As an exception from that pattern, in the *Anticipation* treatment, if a trustee makes a relative return transfer of $Y=1$ and has previously received an investment of $X \geq 20$, he receives on average the same transfer from a helper as a passive player. Furthermore, in the *Anticipation* treatment a relative return transfer of $Y=3$ for a given investment of $X=10$ leads to the same transfer as to player C.

Result 1a: Helpers transfer to trustees, who implemented an equal split in the trust game, does not significantly differ from the transfer to passive players. Relative return transfers that leave the investor with more money than the trustee lead to higher helpers' transfers, while lower return transfers lead to lower helpers' transfers.

V.2. RETURN TRANSFERS

This section analyzes whether trustees correctly anticipate helpers' transfer decisions and therefore reciprocate investors' transfers more in the *Anticipation* treatment than in the *Baseline*. In addition, a positive correlation between investments and return transfers is examined.

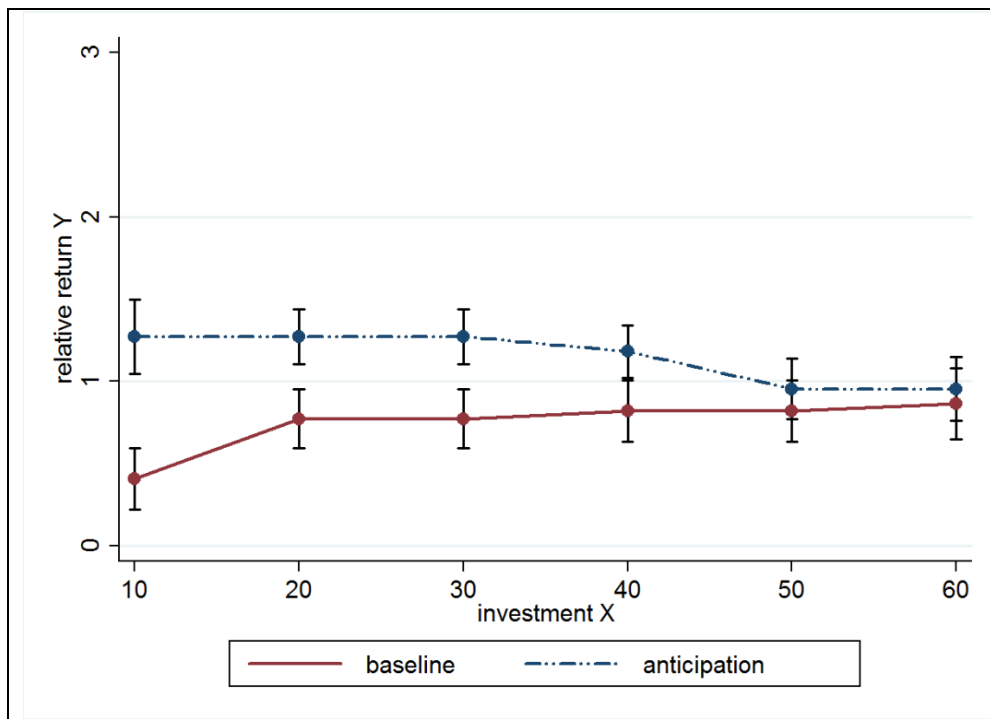


Figure 2

Relative Return Transfers by Treatment

On the x-axis, the investment the trustee has received is depicted; on the y-axis, mean relative return transfers are displayed. Standard errors are indicated.

There are 12 out of 44 trustees who do not make a positive return transfers (9 in *Baseline* and 3 in *Anticipation*).³¹ For the average return transfers, a significant difference between the *Anticipation* treatment and the *Baseline* (Mann-Whitney rank-sum $|z|=2.135$, $p = 0.0327$) is found. In addition, the pattern of the

³¹ The likelihood of any positive return transfer is significantly higher in the *Anticipation* treatment: 1-sided Fisher's exact = 0.044. This is the first indication of the predicted treatment effect by Hypothesis 2b.

results presented in Figure 2 shows a difference between the treatments. Specifically, in the *Baseline*, one can observe a typical outcome for the trust game (see Falk et al. 2011): the higher the investment, the higher the relative return transfer (Spearman's Rho: $r_s=0.1496$, $p=0.0870$). In the *Anticipation* treatment, the reverse occurs: low investments are reciprocated a lot, while the relative returns of high investments do not differ significantly between the treatments (Spearman's Rho: $r_s=-0.1505$, $p=0.0849$).³² The relative return transfers differ significantly between the treatments for investments lower than 50 (Mann-Whitney rank-sum $|z|=1.681$, $p \leq 0.0927$), while they are not statistically different for the highest two possible investments (Mann-Whitney rank-sum $|z|=0.605$, $p \geq 0.5453$).³³ Furthermore, the motives behind the return transfers seem to change in the *Anticipation* treatment compared to the *Baseline*, i.e., subjects "invest" in a reputation of being trustworthy when it is cheap (when a high relative return transfer results in a comparatively low absolute return transfer) and do not reciprocate high investments more than in the *Baseline*. This finding is especially interesting given that studies involving reward by second parties (direct reciprocity) in one-shot, public-good games have not found an increase in socially desirable behavior (e.g., Dufwenberg et al. 2001; Walker and Halloran 2004). Thus, the findings support Hypothesis 2b, while Hypothesis 2a is only supported in the *Baseline*.

Here, a random effects ordered Probit regression is used to test which influence factors determine the return transfers.³⁴ The most important insight from the models in Table 4 is that controlling for different additional influence variables, the data confirm Hypothesis 2b, i.e., the coefficient of the treatment dummy (*Anticipation*) remains highly significant. At first glance, the investment alone (Model 2) does not appear to have a correlation with the return transfer. Yet, one can see a positive significant correlation of the investment and the return transfer when including the interaction of the investment and the treatment dummy (Models 3). The interaction effect is negative and statistically significant. Thus, the results generally support Hypothesis 2a, but the effect is reversed in the *Anticipation* treatment. Summing up, the results of the random effects ordered Probit regression support the findings from the non-parametric data analysis. Overall, the results support Hypothesis 2b, while Hypothesis 2a can only be supported in the *Baseline*. This leads to:

Result 2a: In the *Baseline*, relative return transfers increase as the investments increase. The opposite pattern of the return transfers is found in the *Anticipation* treatment.

³² Comparing for each investment the relative return transfers (return transfer for $X=10$ vs. for $X=20$; for $X=20$ vs. for $X=30$, etc.) overall (and in the *Baseline*), only the first comparison is significantly different ($|z|=1.728$, $p=0.0839$; *Baseline*: ($|z|=2.178$, $p=0.0294$), while the rest is not statistically significantly different ($|z|=1.000$, $p \geq 0.3173$; *Baseline*: ($|z|=1.414$, $p \geq 0.1573$). In the *Anticipation* treatment, only an investment of $X=50$ is reciprocated significantly more than an investment of $X=40$ ($|z|=1.651$, $p=0.0987$), while no other comparison is statistically significantly different ($|z|=1.414$, $p \geq 0.1573$). See Table 9 in the Appendix for all comparisons.

³³ See Table 10 in the Appendix for all non-parametric comparisons.

³⁴ The random effects control for the fact that each trustee takes six decisions (one for each investment) and an ordered Probit model suits these data the best, since trustees can only chose between four different relative return transfer per investment ($Y=0$, $Y=1$, $Y=2$ or $Y=3$).

and to:

Result 2b: The return transfers are, on average, higher in the *Anticipation* treatment than in the *Baseline*.

Table 4
Explaining trustees' transfers – comparison Baseline and Anticipation treatment

| Random effects ordered Probit regressions ("trustees" as group) | | | |
|---|------------------|------------------|------------------|
| Dependent variable: Relative return transfers in the trust game | | | |
| | Model 1 | Model 2 | Model 3 |
| <i>Anticipation</i> | 2.45*** (.27) | 2.45*** (.27) | 4.16*** (.50) |
| Investment | | -.00 (.00) | .02*** (.01) |
| Investment* <i>Anticipation</i> | | | -.04*** (.01) |
| Cut 1 Constant | .29 (.18) | .25 (.25) | 1.13*** (.34) |
| Cut 2 Constant | 2.29*** (.21) | 2.25*** (.27) | 3.28*** (.38) |
| Cut 3 Constant | 4.54*** (.39) | 4.50*** (.41) | 5.65*** (.52) |
| Rho Constant | .86*** (.02) | .86*** (.02) | .87*** (.02) |
| N | 264 | 264 | 264 |
| N of groups | 44 | 44 | 44 |
| P model | 0.033 | 0.101 | <.001 |
| LR Chi2 | 4.52 | 4.58 | 23.73 |

Random effects ordered Probit regressions. Standard errors are presented in parentheses. The *Anticipation* dummy equals 1 for all observations of the *Anticipation* treatment, *investment* controls for the investment (X) the trustee has received. Significance at the 10%, 5%, and 1% level is denoted by *, **, and ***, respectively.

V.3. INVESTMENTS

This section tests whether investors correctly anticipate the different return transfers between the treatments and therefore invest more in the *Anticipation* treatment compared to the *Baseline*.

Statistically, the investments do not, on average, differ between the treatments (Mann-Whitney rank-sum: $|z| = 0.465$, $p = 0.6419$), but the cumulative distribution function of the investments in Figure 3 shows that the distribution of the investments differs. Specifically, in the *Anticipation* treatment, zero-

investments are never made, whereas in the *Baseline*, zero-transfers are present.³⁵ Nonetheless, the data do not support Hypothesis 3 and lead to:

Result 3: On average, investments do not statistically differ significantly between treatments.

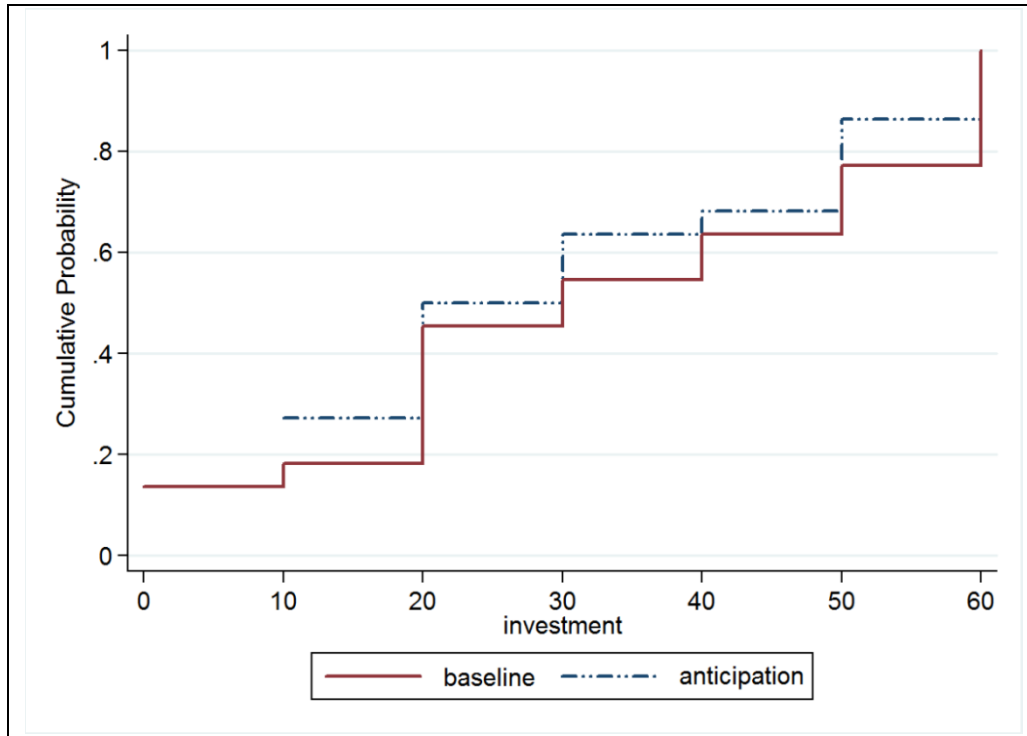


Figure 3

Investments by Treatment

The graph shows an empirical cumulative distribution function of the investments for both treatments. On the x-axis, the investment is depicted; on the y-axis, the estimated probability for each investment is displayed.

V.4. EARNINGS

As mentioned above, helpers seem to care more about socially desirable outcomes than about the motives behind trustees' transfers. Those preferences raise the question of whether the monetary efficiency is, in fact, higher in the *Anticipation* treatment. As shown above, helpers make, on average, the same transfers in the *Anticipation* treatment as in the *Baseline* given a particular return transfer and trustees anticipate that. I.e., trustees make higher relative return transfers in the *Anticipation* treatment especially if such an

³⁵ Using Fisher's exact test for a positive investment does not lead to a statistically significant treatment difference (1-sided Fisher's exact = 0.116). Testing the data parametrically with an ordered Probit regression model does not lead to any significant results either.

"investment" in a good reputation is cheap in absolute terms. It turns out that the money-maximizing strategy for trustees (given the strategy of the helpers) in the *Anticipation*³⁶ treatment is fully to return ($Y=3$) each investment (X) (see Table 11 in the Appendix for a calculation of possible payoffs for each strategy of a trustee). Thus, trustees do not anticipate helpers' positive strong indirect reciprocity enough to maximize their profits.

Given the decisions of trustees investors' money-maximizing strategy is to invest $X=0$ in *Baseline* and $X=30$ in the *Anticipation* treatment (see Table 12 in the Appendix for a calculation of possible payoffs for each investment of an investor). However, the quantitative difference in payoffs is very small (100 compared to 108.18). If investors are at least slightly risk-averse, they make the correct decision by not investing differently in the *Anticipation* treatment compared to the *Baseline*. An alternative explanation for the underinvestment could be that the level of reasoning is too high. I.e., investors have to anticipate not only what trustees will do, but also what they think what trustees think the helpers will do in the later game.

37

Now, the actually realized transfers are presented. Since average realized return transfers in the trust game are statistically significantly higher in the *Anticipation* treatment than in the *Baseline* (*Baseline*: 26.36 (sd=34.98); *Anticipation*: 41.81 (sd=34.03); Mann-Whitney rank-sum $|z|= 1.906$, $p = 0.0566$) realized helpers' transfers are (descriptively) higher in the *Anticipation* treatment than in the *Baseline* (*Baseline*: 17.72 (sd=29.86); *Anticipation*: 28.63 (sd=50.26); Mann-Whitney rank-sum: $|z| = 0.404$, $p = 0.6861$). As a result, Players B (trustees in the trust game and helpers' co-players in the helping game) have, on average, the same total earning in the *Anticipation* treatment as in the *Baseline* (*Baseline*: 188.18 (sd=55.02); *Anticipation*: 173.63 (sd=75.15); Mann-Whitney rank-sum: $|z| = 1.506$, $p = 0.1321$).³⁸ In total, players A have, on average, the same profits in the *Baseline* and in the treatment (*Baseline*: 187.72 (sd=30.96); *Anticipation*: 197.27 (sd=32.61); Mann-Whitney rank-sum: $|z| = 0.969$, $p = 0.3324$). This finding, however, should be treated with caution, since the earnings of the players arise from the particular design and the matching in the experiment, i.e., trustees become helpers' co-players in the helping game not only in the *Anticipation* treatment, but also in the *Baseline*, and Players A also have the role of investors as of a helpers.

³⁶ Obviously, in the *Baseline* (as discussed in chapter IV), in expectations the money-maximizing strategy for trustees is to send zero transfers (since in the *Baseline* trustees do not know that the helping game will follow). Similarly, (as well discussed in chapter IV) the money-maximizing strategy in both treatments for helpers is to send zero transfer. Quantitatively, trustees follow the optimal strategy (see Table 12 in the Appendix).

³⁷ See the level-k literature, as in Nagel (1995), Stahl et al. (1995), Ho et al. (1998), Costa-Gomes et al. (2001), and Costa-Gomes et al. (2006), etc.

³⁸ Here, only trustees who become co-players of helpers in the helping game are included. The result does not change if trustees who do not become helpers' co-players in the second part of the experiment are included. The total earnings for the C-players do not differ significantly between the treatments either, irrespectively of whether they become co-players of helpers (Mann-Whitney rank-sum: $|z| = 0.408$, $p = 0.6831$) or not (Mann-Whitney rank-sum: $|z| = 0.298$, $p = 0.7660$).

VI. CONCLUSION

From a welfare point of view, rewards are better for the society than punishment since they do not lead to an efficiency loss. Rand et al. (2009) show in a repeated public goods game that "*reward is as effective as punishment for maintaining public cooperation and leads to higher total earnings. Moreover, when both options are available, reward leads to increased contributions and payoff, whereas punishment has no effect on contributions and leads to lower payoff*" (Rand et al. 2009, p. 1272). In a world in which anonymous interactions become more and more frequent (e.g., via the internet), it is important to pay more attention to economic consequences of reward systems by impartial strangers. Surprisingly, until now only a small number of papers analyzes rewards rather than punishments and even less literature considers rewards given by non-strategically motivated third parties. Most notably, there exists no literature on the question: How can the anticipation of a reward from an impartial stranger enhance trustworthiness (and cooperation)? The current paper attempts to close this gap.

This paper shows that positive strong indirect reciprocity exists and, moreover, that it is correctly anticipated by potential recipients. Helpers' transfers are surprisingly more positively correlated with relative return transfers in the *Anticipation* treatment than in the *Baseline*. Trustees anticipate this behavior correctly and make higher return transfers in the *Anticipation* treatment. In particular, they make higher relative return transfers in response to lower investments, signaling higher trustworthiness when such signals are cheap in absolute terms. Despite the fact that the earnings of players do not differ between the treatments, trustees become more trustworthy in the *Anticipation* treatment than in the *Baseline* and they would have earned even more money than their counterparts in the *Baseline* if they had fully anticipated the complete strategy of helpers.

The experiment still leaves some questions open. There is high zero-censoring in the helping game. This could be interpreted as helpers' willingness to punish greedy trustees. An experiment where real punishment is used instead of denying help may be used to understand the helpers' actions further. Dohmen et al. (2009) show that positive and negative reciprocity are different concepts. This finding might be true for strong negative and positive indirect reciprocity as well. However, without encountering ethical concerns, one could implement only the announced version of the punishment game.

This paper helps to explore the motivations behind and the consequences of third-party reward of trustworthiness. The findings from this paper are applicable to many different settings. The results show that, in all situations in which people display trustworthy behavior for intrinsic reasons, they might in fact do a better job if they could expect strangers to reward them for their actions. That is even true if the strangers do not have any strategic incentives to do so and if their actions are costly. From a policy perspective, it turns out that it can be socially beneficial to promote greater publicity on socially desirable acts.

VII. REFERENCES

- ABELER, JOHANNES, ARMIN FALK, LORENZ GOETTE and DAVID HUFFMAN (2011). "Reference points and effort provision." American Economic Review **101**(2): 470–492.
- ALMENBERG, JOHAN, ANNA DREBER, COREN L. APICELLA and DAVID G. RAND (2011). "Third Party Reward and Punishment: Group Size, Efficiency and Public Goods." In Psychology and Punishment (Palmetti, N.M. and Russo, J.P., eds): 73–92, Nova Science Publishers.
- ANDREONI, JAMES (1990). "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." The Economic Journal **100**(401): 464–477.
- ANDREONI, JAMES and JOHN H. MILLER (1993) "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence" The Economic Journal **103**(418): 570-585.
- ARROW, KENNETH. J. (1972). "Gifts and Exchanges" Philosophy and Public Affairs **1**(4): 343-362.
- BALAFOUTAS, LUCAS, NICOS NIKIFORAKIS and KRISTOFFEL GRECHENIG (2014) "Third-party punishment and counter-punishment in one-shot interactions" Economics Letters **122**(2): 308–310.
- BERG, JOYCE, JOHN DICKHAUT and KEVIN MCCABE (1995). "Trust, reciprocity, and social history." Games and Economic Behavior **10**: 122–142.
- BERNHARD, HELEN, URS FISCHBACHER and ERNST FEHR (2006). "Parochial altruism in humans" Nature **442**: 912-915.
- BRANDTS, JORDY and GARY CHARNES (2011). "The strategy versus the direct-response method: A first survey of experimental comparisons." Experimental Economics **14**(3): 375–398.
- BARDSLEY, NICOLAS (2008). "Dictator Game Giving: Altruism or Artefact?" Experimental Economics **11**(2): 122-133.
- BOYD, ROBERT and PETER J. RICHESON (1989). "The evolution of indirect reciprocity." Social Networks **11**: 213–236.
- CAMERER, COLIN F. and ERNST FEHR (2004) "Measuring Social Norms and Preferences Using Experimental Games: A Guide for Social Scientists", in HENRICH, JOSEPH, ROBERT BOYD, SAMUEL BOWLES, COLIN CAMERER, ERNST FEHR and HERBERT GINTIS eds. (2004) Foundations of human sociality: Economic experiments and ethnographic evidence from fifteen small-scale societies. New York: Oxford University Press.

- CARPENTER, JEFFREY P. and PETER HANS MATTHEWS (2004). "Social reciprocity." IZA Discussion Paper No. **1347**.
- CARPENTER, JEFFREY P., PETER HANS MATTHEWS and OKOMBOLI ONG'ONG'A (2004). "Why Punish? Social reciprocity and the enforcement of prosocial norms." Journal of Evolutionary Economics **14**: 407–429.
- CHARNESS, GARY and MATTHEW RABIN (2002). "Understanding social preferences with simple tests." The Quarterly Journal of Economics **117**(3): 817-869.
- COSTA-GOMES, MIGUEL and VINCENT P. CRAWFORD (2006). "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study." American Economic Review **96**(5): 1737-1768.
- COSTA-GOMES, MIGUEL, VINCENT P. CRAWFORD and BRUNO BROSETA (2001). "Cognition and Behavior in Normal-Form Games: An Experimental Study." Econometrica **69**(5): 1193-1235.
- COSTA-GOMES, MIGUEL, STEFFEN HUCK and GEORG WEIZSÄCKERBERLIN (2012). "Beliefs and actions in the trust game: Creating instrumental" WZB Discussion Papers **SP II 2012-302**.
- DOHMEN, THOMAS, ARMIN FALK, DAVID HUFFMAN and UWE SUNDE (2009). "Homo Reciprocans: Survey Evidence on Behavioural Outcomes." The Economic Journal **119**(536): 592–612.
- DUFWENBERG, MARTIN, URI GNEEZY, WERNER GÜTH and ERIC VAN DAMME (2001). "Direct versus indirect reciprocity: An experiment." Homo Oeconomicus **18**: 19–30.
- DUFWENBERG, MARTIN and GEORG KIRCHSTEIGER (2004). "A theory of sequential reciprocity." Games and Economic Behavior **47**: 268–298.
- ENGELMANN, DIRK and URS FISCHBACHER (2009). "Indirect reciprocity and strategic reputation building in an experimental helping game." Games and Economic Behavior **67**: 399–407.
- FALK, ARMIN and URS FISCHBACHER (2006). "A Theory of Reciprocity." Games and Economic Behavior **54**(2): 293-315.
- FALK, ARMIN, ERNST FEHR and URS FISCHBACHER (2008). "Testing theories of fairness – Intentions matter." Games and Economic Behavior **62**(1): 287–303.
- FALK, ARMIN, STEPHAN MEIER and CHRISTIAN ZEHNDER (2011). "Do Lab Experiments Misrepresent Social Preferences? The Case of Self-Selected Student Samples" Journal of the European Economic Association in press.

- FEHR, ERNST and SIMON GÄCHTER (2000). "Cooperation and punishment in public goods experiments." American Economic Review **90**: 980-994.
- FEHR, ERNST and URS FISCHBACHER SIMON GÄCHTER (2004). "Third-party punishment and social norms." Evolution and Human Behavior **25**: 63 – 87.
- FEHR, ERNST, URS FISCHBACHER and SIMON GÄCHTER (2002). "Strong reciprocity, human cooperation and the enforcement of social norms." Human Nature **13**: 1–25.
- FISCHBACHER, URS (2007). "z-Tree. Zurich Toolbox for Ready-made Economic Experiments." Experimental Economics **10**(2): 171-178.
- FISCHBACHER, URS, SIMON GÄCHTER and SIMONE QUERCIA (2012). "The Behavioural Validity of the Strategy Method in Public Good Experiments." Journal of Economic Psychology **33**(4): 897-913.
- FORSYTHE, ROBERT, JOEL L. HOROWITZ, NATHAN E. SAVIN and MARTIN SEFTON (1994). "Fairness in Simple Bargaining Games" Games and Economic Behavior **6**(3): 347–369.
- GREINER, BEN (2004). "An Online Recruitment System for Economic Experiments" in Forschung und Wissenschaftliches Rechnen 2003, ed. by K. Kremer and V. Macho. GWDG Bericht 63. Gesellschaft für Wissenschaftliche Datenverarbeitung Göttingen: Datenverarbeitung, 79–93.
- GREINER, BEN and M. VITTORIA LEVATI (2005). "Indirect reciprocity in cyclical networks. An experimental study." Journal of Economic Psychology **26**: 711–731.
- GÜTH, WERNER, MANFRED KONIGSTEIN, NADÈGE MARCHAND and KLAUS NEHRING (2001). "Trust and reciprocity in the investment game with indirect reward." Homo Oeconomicus **18**: 241–262.
- HOLT, CHARLES A. and SUSAN K. LAURY (2002). "Risk Aversion and Incentive Effects." American Economic Review **92**(5): 1644-55.
- LEVINE, DAVID K. (1998). "Modeling Altruism and Spitefulness in Experiments" Review of Economic Dynamics **1**: 593-622.
- NAGEL, ROSEMARIE (1995). "Unraveling in Guessing Games: An Experimental Study." American Economic Review **85**(5): 1313–1326.
- RABIN, MATTHEW (1993). "Incorporating fairness into game theory and economics." American Economic Review **83**: 1281–1302.
- RAND, DAVID G., ANNA DREBER, TORE ELLINGSEN, DREW FUDENBERG and MARTIN A. NOWAK (2009). "Positive Interactions Promote Public Cooperation" Science **325**: 1272-1275.

- ROCKENBACH, BETTINA and MANFRED MILINSKI (2006). "The efficient interaction of indirect reciprocity and costly punishment." Nature **444**: 718–723.
- SEINEN, INGRID and ARTHUR SCHRAM (2006). "Social status and group norms: Indirect reciprocity in a repeated helping experiment." European Economic Review **50**: 581–602.
- SELTEN, REINHARD (1967). "Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments." Beiträge zur experimentellen Wirtschaftsforschung. E. Sauer mann. Tübingen, Mohr: 136-168.
- STANCA, LUCA (2009). "Measuring Indirect Reciprocity: Whose Back Do We Scratch?" Journal of Economic Psychology **30**(2): 190-202.
- STANCA, LUCA, LUIGINO BRUNI and LUCA CORAZZINI (2009). "Testing Theories of Reciprocity: Do Motivations Matter?" Journal of Economic Behavior and Organization **71**(2): 233-245.
- WALKER, JAMES M. and MATTHEW HALLORAN (2004). "Rewards and Sanctions and the Provision of Public Goods in One-Shot Settings." Experimental Economics **7**(3): 235-247.

VIII. APPENDIX

VIII.1 INSTRUCTIONS

Here, the English translation of the instructions is presented. The original German version can be passed on upon request. The only difference between the *Anticipation* treatment and the *Baseline* is that, in the *Anticipation* treatment, first only the general instructions and the instructions for the trust game are presented and only after the trust game is played, the instructions for the helping game are handed out. On the contrary, in the *Baseline* the full instructions are handed out immediately at the beginning of the experiment. The instructions and the control questions are presented here in the order in which they are read out aloud by the experimenter (and filled out by the participants).

| |
|---------------------------------------|
| General Instructions for Participants |
|---------------------------------------|

You are about to take part in economic experiments. Depending on the decisions that you and others make, you can earn a substantial amount of money. It is therefore very important that you read these instructions carefully.

The written statements you have received from us serve your own private information only. **During the experiments, any communication whatsoever is forbidden.** If you have any questions, please ask only us. Please raise your hand and we will come to you. Disobeying these rules will lead to exclusion from the experiments and from all payments.

During the experiments, we speak not of Euro, but instead of Taler. Your entire income is hence initially calculated in Taler. The total number of Taler you earn during the experiments is converted into Euro at the end and paid to you in cash, at the rate of

1 Taler = 2 Eurocent.

In addition, each participant is paid a lump sum of 4 Euro for showing up today.

You will take part in several experiments today. The instructions to each experiment will be handed out to you one by one, just before the respective experiment is about to begin. On the following pages, we will describe the exact procedure of the first experiment.

In this experiment, there are three different roles: Player A, Player B, and Player C. At the beginning of the experiment, you are randomly assigned one of these three roles.

In the experiment, you are required to make your decision once only; i.e., the experiment is conducted only once. You will thus make no repeated decisions today.

Information on the **First Experiment**

In this experiment, one Player A and one Player B are randomly paired. Each player initially receives 100 Taler, which we shall refer to in the following as the initial endowment.

This experiment consists of 2 stages.

Stage 1: In the first stage, **Player A** decides which transfer of X Taler from the initial endowment he or she wishes to send to B. X may be one of the following values: 0, 10, 20, 30, 40, 50, or 60. The Taler sent by Player A are **tripled** and transferred to the Player B who has been assigned to him/her.

Stage 2: If Player A has sent more than 0 Taler in Stage A, **Player B** now has to decide, in Stage 2, how many Taler $Y \cdot X$ he or she wishes to transfer back to A. Player B may choose from the following options: no Taler ($Y = 0$), the same amount of Taler as in the transfer ($Y = 1$), double the transferred amount of Taler ($Y = 2$), or three times the transferred amount of Taler ($Y = 3$).

The income at the end of the first part of the experiment is therefore:

Player A:

If he/she has sent 0: 100

If he/she has sent X: $100 - X + (X * Y)$

Player B:

If he/she receives 0: 100

If he/she receives X: $100 + 3 * X - (X * Y)$

Here, X may be equal to 0, to 10, to 20, to 30, to 40, to 50, or to 60.

Y may be equal to 0, to 1, to 2, or to 3.

Examples:

1.) If, for example, Player A sends a sum of $X = 10$ Taler and B decides to send back three times the amount of the transfer ($Y = 3$), then both players have the following income:

A: $100 - 10 + (10 * 3) = 120$, and B: $100 + 3 * 10 - 3 * 10 = 100$.

2.) If, for example, Player A sends a sum of $X = 60$ Taler and B decides to send back the amount of the transfer ($Y = 1$), then both players have the following income:

A: $100 - 60 + (60 * 1) = 100$, and B: $100 + 3 * 60 - 1 * 60 = 220$.

3.) If, for example, Player A sends a sum of $X = 40$ Taler and B decides to send back twice the amount of the transfer ($Y = 2$), then both players have the following income:

A: $100 - 40 + (40 * 2) = 140$, and B: $100 + 3 * 40 - 2 * 40 = 140$.

However, only at the end of Stage 2 are you told how high your income is from the experiment.

Player C does not have to make any decision in this experiment. Player C's payoff in this experiment is 100 Taler.

The experiment ends here. You are then told how high your income is from the experiment. Further experiments follow – **however, it is impossible for you to be assigned once again to a group with the same players. Further, you cannot lose your payoffs from the experiments.** Following the final experiment, you will be given a questionnaire. Once you have filled in the questionnaire, you will receive your payoff from us in cash. In order to receive your payoff, please bring all documents you have received from us with you.

| |
|--------------------------------------|
| Control Questions for Experiment One |
|--------------------------------------|

1.) Assuming that Player A sends 50 Taler to Player B and Player B sends 50 Taler back to Player A,

a. How high is Player A's income? _____

b. How high is Player B's income? _____

2.) Assuming that Player A sends 20 Taler to Player B and Player B sends 60 Taler back to Player A,

a. How high is Player A's income? _____

b. How high is Player B's income? _____

3.) Assuming that Player A sends 0 Taler to Player B,

a. Can Player B send something back to Player A? _____

b. How high is Player A's income? _____

c. How high is Player B's income? _____

4.) How high is Player C's income? _____

Information on the **Second Experiment**

Participants in this experiment have the same roles as in the first experiment. Hence, this means that a participant who had role A in the first experiment will still have role A in this experiment; a participant who had role B in the first experiment will still have role B in this experiment; and a participant who had role C in the first experiment will still have role C in this experiment. Random assignation is conducted anonymously, and we ensure that **you are not assigned once again to a group with the same players as in the first experiment.**

This second experiment also consists of two stages.

Stage 1: In the first stage, **Player A** receives an endowment of 100 Taler. He or she now has to decide how many of these 100 Taler to send his or her co-player (full numbers between 0 and 100). Every Taler sent is **tripled** and credited to the other player's account.

Whether the co-player is a Player B or a Player C is decided randomly only in Stage 2. However, Player A must make a binding decision in Stage 1 on how many Taler to send if the co-player is a Player B, and how many Taler to send if the co-player is a Player C. In addition, if it is a Player B, Player A may also decide how many Taler to send depending on Player B's behavior in the first experiment. In Stage 1, Player A therefore has to fill in the following seven decision tables for all possible transfers X and $Y * X$:

This is what the decision tables look like:

1st Screen

If, in Stage 2, my co-player is ...

- ... a **Player C**, ...
... I will now send _____ Taler.
- ... a **Player B** who has been sent **0 Taler** in the first experiment, ...
... I will now send _____ Taler.

2nd Screen

If, in Stage 2, my co-player is ...

- ... a **Player B** who has been sent **10 Taler** in the first experiment and ...
 - ... sent back **0 Taler**, I will now send ____ Taler.
 - ... sent back **10 Taler**, I will now send ____ Taler.
 - ... sent back **20 Taler**, I will now send ____ Taler.
 - ... sent back **30 Taler**, I will now send ____ Taler.

3rd Screen

If, in Stage 2, my co-player is ...

- ... a **Player B** who has been sent **20 Taler** in the first experiment and ...
 - ... sent back **0 Taler**, I will now send ____ Taler.
 - ... sent back **20 Taler**, I will now send ____ Taler.
 - ... sent back **40 Taler**, I will now send ____ Taler.
 - ... sent back **60 Taler**, I will now send ____ Taler.

4th Screen

If, in Stage 2, my co-player is ...

- ... a **Player B** who has been sent **30 Taler** in the first experiment and ...
 - ... sent back **0 Taler**, I will now send ____ Taler.
 - ... sent back **30 Taler**, I will now send ____ Taler.
 - ... sent back **60 Taler**, I will now send ____ Taler.
 - ... sent back **90 Taler**, I will now send ____ Taler.

5th Screen

If, in Stage 2, my co-player is ...

- ... a **Player B** who has been sent **40 Taler** in the first experiment and ...
... sent back **0 Taler**, I will now send ____ Taler.
... sent back **40 Taler**, I will now send ____ Taler.
... sent back **80 Taler**, I will now send ____ Taler.
... sent back **120 Taler**, I will now send ____ Taler.

6th Screen

If, in Stage 2, my co-player is ...

- ... a **Player B** who has been sent **50 Taler** in the first experiment and ...
... sent back **0 Taler**, I will now send ____ Taler.
... sent back **50 Taler**, I will now send ____ Taler.
... sent back **100 Taler**, I will now send ____ Taler.
... sent back **150 Taler**, I will now send ____ Taler.

7th Screen

If, in Stage 2, my co-player is ...

- ... a **Player B** who has been sent **60 Taler** in the first experiment and ...
... sent back **0 Taler**, I will now send ____ Taler.
... sent back **60 Taler**, I will now send ____ Taler.
... sent back **120 Taler**, I will now send ____ Taler.
... sent back **180 Taler**, I will now send ____ Taler.

You may enter any number between 0 and 100 in each line. It goes without saying that the lines do **not** have to add up to 100 either, for **only that line is decision-relevant which actually corresponds to the situation drawn by lot in Stage 2!** The decisions in the other lines (not drawn) do not influence your payoff. Do please note, however, that while filling in the table you do not yet

know if your co-player is a Player C or B (and you do not know, in case a Player B is assigned to you, how the he has behaved in the first part of the experiment). **In each line, you therefore have to consider your decision carefully, for every one can become relevant for you.**

Stage 2: In the second stage, it is decided by draw whether Player A's co-player is a Player B or C. The player who is drawn then receives the amount of Taler according to the corresponding decision table from Stage 1. He or she therefore does not have to make any decision in this second experiment.

The other player – who is not drawn – is given the possibility to increase his or her own income individually, by way of a small task. (Precise instructions for this task will appear later on this player's screen.)

The income from the second experiment is therefore:

Player A:

$100 - (\text{what A sent the co-player})$

Co-player (B or C, depending on the draw in Stage 2):

$3 * (\text{what A sent the co-player})$

Player not drawn (B or C, depending on the draw in Stage 2):

Income from the individual small task

Examples:

1.) If Player A should hence decide to enter the following numbers in the second table: top line 1, second line 13, third line 17, bottom line 0; and if the co-player assigned to Player A is a Player B who was sent 10 Taler by his co-player in the first experiment, and who in turn decided to send nothing back, then the payoffs from this second experiment are as follows for Player A, who made the decision here:

$100 - 1 = 9$. The payoff for the co-player assigned to Player A is: $1 * 3 = 3$.

2.) If Player A should decide to enter the following numbers in the first table: top line 99, second line 14; and if the co-player assigned to Player A is a Player C, then the payoffs from this second experiment are as follows for Player A, who has made the decision here:

$100 - 99 = 1$. The payoff for the co-player assigned to Player A is: $99 * 3 = 297$.

The experiment ends here. You will then hear about your payoff from the first and second experiment. Further experiments will follow – **however, it is impossible for you to be assigned once again to a group with the same players. Further, you cannot lose your payoffs from the experiments.** Following the final experiment, you will be given a questionnaire. Once you have filled in the questionnaire, you will receive your payoff from us in cash. In order to receive your payoff, please bring all documents you have received from us with you.

Control Questions for Experiment Two

- 1.) We assume that Player A in Experiment 2 is assigned to a Player B who was sent 10 Taler by Player A in Experiment 1 and sent back 30 Taler to Player A. If Player A fills in the table in Experiment Two in the following way:

If, in Stage 2, my co-player is...

- ... a **Player B** who was sent **10 Taler** in the first experiment and ...
 - ... sent back **0 Taler**, I will now send __5__Taler.
 - ... sent back **10 Taler**, I will now send __17__Taler.
 - ... sent back **20 Taler**, I will now send __0__Taler.
 - ... sent back **30 Taler**, I will now send __21__Taler.

- a. How high is Player A's income from Experiment 2? _____
- b. How high is Player B's income from Experiment 2? _____

- 2.) We assume that Player A in Experiment 2 is assigned to a Player C. If Player A fills in the table in Experiment Two in the following way:

If, in Stage 2, my co-player is...

- ... a **Player C**,
 - ... I will now send __50__Taler.
- ... a **Player B** who was sent **0 Taler** in the first experiment,...
 - ... I will now send __50__Taler.

- a. How high is Player A's income from Experiment 2? _____
- b. How high is Player C's income from Experiment 2? _____

Information on the **Third Experiment**

In this part of the experiment, no other participant is paired with you. The payoffs therefore relate only to you. The decisions of the other participants only have an influence on their own respective payoffs.

In this part of the experiment, you are asked to decide in 10 different situations (lotteries) between option A and B. These situations will be presented to you on consecutive screens. The two lotteries each comprise 2 possible monetary payoffs, one high and one low, which will be paid to you with different probabilities.

The options A and B will be presented to you on the screen, as in the following example:

Lottery 1
Please choose the lottery you prefer.

| Lottery A: | | |
|-------------|--------|--------|
| Probability | 1/10 | 9/10 |
| Payoff | 2.00 € | 1.60 € |

A

| Lottery B: | | |
|-------------|--------|--------|
| Probability | 1/10 | 9/10 |
| Payoff | 3.85 € | 0.10 € |

B

The computer uses a random draw program, which assigns you payments exactly according to the denoted probabilities.

For the above example, this means:

Option A obtains a payoff of 2 Euro with a probability of 10% and a payoff of 1.60 Euro with a probability of 90%.

Option B obtains a payoff of 3.85 Euro with a probability of 10% and a payoff of 0.10 Euro with a probability of 90%.

Now you have to click on the particular option you decide for.

Please note that, at the end of the experiment, only one of the 10 situations will eventually be paid. Yet, each of the situations can be randomly chosen with equal probability to be the payoff-relevant one.

After this, a draw will determine whether for the payoff-relevant situation the high payoff (2.00 Euro or 3.85 Euro) or the low payoff (1.60 Euro or 0.10 Euro) will be paid.

VIII.2 TABLES

TABLE 5

Pairwise comparison of transfers to active Player B in the helping game per increased return transfer by one for any given possible investment

| Wilcoxon signed-rank tests: p-values, z in parentheses | | | | |
|---|----------------------------|------------------------------------|-----------------|---------------------|
| <i>Investment (X)</i> | <i>Return Transfer (Y)</i> | <i>Baseline & Anticipation</i> | <i>Baseline</i> | <i>Anticipation</i> |
| X = 10 | Y = 0 vs. Y = 1 | 0.0008, (3.346) | 0.0050, (2.810) | 0.0546, (1.922) |
| X = 10 | Y = 1 vs. Y = 2 | 0.0000, (4.886) | 0.0007, (3.398) | 0.0004, (3.515) |
| X = 10 | Y = 2 vs. Y = 3 | 0.0002, (3.753) | 0.0125, (2.498) | 0.0049, (2.811) |
| X = 20 | Y = 0 vs. Y = 1 | 0.0009, (3.315) | 0.0627, (1.861) | 0.0050, (2.808) |
| X = 20 | Y = 1 vs. Y = 2 | 0.0000, (4.434) | 0.0025, (3.019) | 0.0012, (3.251) |
| X = 20 | Y = 2 vs. Y = 3 | 0.0007, (3.409) | 0.0642, (1.851) | 0.0034, (2.928) |
| X = 30 | Y = 0 vs. Y = 1 | 0.0000, (4.087) | 0.0084, (2.636) | 0.0018, (3.122) |
| X = 30 | Y = 1 vs. Y = 2 | 0.0000, (4.581) | 0.0007, (3.396) | 0.0021, (3.080) |
| X = 30 | Y = 2 vs. Y = 3 | 0.0000, (4.473) | 0.0125, (2.496) | 0.0002, (3.733) |
| X = 40 | Y = 0 vs. Y = 1 | 0.0000, (4.087) | 0.0085, (2.633) | 0.0018, (3.123) |
| X = 40 | Y = 1 vs. Y = 2 | 0.0000, (4.966) | 0.0007, (3.394) | 0.0003, (3.625) |
| X = 40 | Y = 2 vs. Y = 3 | 0.0000, (4.131) | 0.0750, (1.781) | 0.0001, (3.908) |
| X = 50 | Y = 0 vs. Y = 1 | 0.0001, (3.847) | 0.0146, (2.441) | 0.0030, (2.971) |
| X = 50 | Y = 1 vs. Y = 2 | 0.0000, (4.748) | 0.0037, (2.906) | 0.0002, (3.726) |
| X = 50 | Y = 2 vs. Y = 3 | 0.0000, (4.376) | 0.0276, (2.204) | 0.0001, (3.824) |
| X = 60 | Y = 0 vs. Y = 1 | 0.0025, (3.018) | 0.0600, (1.881) | 0.0171, (2.384) |
| X = 60 | Y = 1 vs. Y = 2 | 0.0000, (4.591) | 0.0052, (2.794) | 0.0003, (3.627) |
| X = 60 | Y = 2 vs. Y = 3 | 0.0003, (3.622) | 0.0864, (1.715) | 0.0014, (3.202) |
| The shaded area highlights all comparisons that do not lead to a significant effect at the 10-percent level. | | | | |

TABLE 6

Pairwise comparison of transfers to active Player B vs. to passive players in the helping game for any possible investment and return transfer

| Wilcoxon signed-rank tests: p-values, z in parentheses | | | | | | | |
|--|--------------------------------|---|-----------------|---------------------|---|-----------------|---------------------|
| <i>Investment (X)</i> | <i>Return Transfer (Y)</i> | <i>To active Player B vs. to passive Player B (X=0)</i> | | | <i>To active Player B vs. to Player C</i> | | |
| | | <i>Baseline & Anticipation</i> | <i>Baseline</i> | <i>Anticipation</i> | <i>Baseline & Anticipation</i> | <i>Baseline</i> | <i>Anticipation</i> |
| X = 10 | Y = 0 | 0.0000, (4.282) | 0.0080, (2.653) | 0.0050, (2.809) | 0.0001, (4.041) | 0.0018, (3.121) | 0.0072, (2.686) |
| X = 20 | Y = 0 | 0.0000, (4.270) | 0.0061, (2.742) | 0.0052, (2.792) | 0.0000, (4.085) | 0.0030, (2.970) | 0.0053, (2.787) |
| X = 30 | Y = 0 | 0.0000, (4.425) | 0.0011, (3.266) | 0.0026, (3.015) | 0.0000, (4.611) | 0.0018, (3.122) | 0.0007, (3.392) |
| X = 40 | Y = 0 | 0.0000, (4.394) | 0.0011, (3.264) | 0.0032, (2.946) | 0.0000, (4.611) | 0.0030, (2.970) | 0.0004, (3.514) |
| X = 50 | Y = 0 | 0.0000, (4.066) | 0.0011, (3.266) | 0.0026, (3.015) | 0.0000, (4.148) | 0.0132, (2.480) | 0.0007, (3.392) |
| X = 60 | Y = 0 | 0.0001, (3.885) | 0.0011, (3.263) | 0.0045, (2.843) | 0.0001, (3.911) | 0.0161, (2.408) | 0.0015, (3.176) |
| X = 10 | Y = 1 | 0.0053, (2.791) | 0.0168, (2.391) | 0.0906, (1.692) | 0.0029, (2.974) | 0.0748, (1.781) | 0.0170, (2.387) |
| X = 20 | Y = 1 | 0.0058, (2.760) | 0.0044, (2.847) | 0.2203, (1.226) | 0.0032, (2.946) | 0.0029, (2.975) | 0.1655, (1.387) |
| X = 30 | Y = 1 | 0.0081, (2.649) | 0.0044, (2.850) | 0.3075, (1.020) | 0.0130, (2.484) | 0.0050, (2.810) | 0.3412, (0.952) |
| X = 40 | Y = 1 | 0.0358, (2.099) | 0.0230, (2.273) | 0.3892, (0.861) | 0.0607, (1.875) | 0.0252, (2.238) | 0.4972, (0.679) |
| X = 50 | Y = 1 | 0.0529, (1.935) | 0.0230, (2.273) | 0.5367, (0.618) | 0.0679, (1.825) | 0.0252, (2.238) | 0.5403, (0.612) |
| X = 60 | Y = 1 | 0.0698, (1.813) | 0.0239, (2.259) | 0.5703, (0.568) | 0.0185, (2.356) | 0.0264, (2.220) | 0.2357, (1.186) |
| X = 10 | Y = 2 | 0.3227, (0.989) | 0.3785, (0.881) | 0.5543, (0.591) | 0.5473, (0.602) | 0.5087, (0.661) | 0.8261, (0.220) |
| X = 20 | Y = 2 | 0.7258, (0.351) | 0.2242, (1.215) | 0.6396, (0.468) | 0.5026, (0.670) | 0.9546, (0.057) | 0.3614, (0.913) |
| X = 30 | Y = 2 | 0.8972, (0.129) | 0.6173, (0.500) | 0.8150, (0.234) | 0.4039, (0.835) | 0.7412, (0.330) | 0.3911, (0.858) |
| X = 40 | Y = 2 | 0.9013, (0.124) | 0.3887, (0.862) | 0.4521, (0.752) | 0.4279, (0.793) | 0.9094, (0.114) | 0.3287, (0.977) |
| X = 50 | Y = 2 | 0.7783, (0.281) | 0.4479, (0.759) | 0.3673, (0.902) | 0.1415, (1.470) | 0.4864, (0.696) | 0.1601, (1.405) |
| X = 60 | Y = 2 | 0.4420, (0.769) | 0.7548, (0.312) | 0.2798, (1.081) | 0.1494, (1.442) | 0.3825, (0.873) | 0.2177, (1.233) |
| X = 10 | Y = 3 | 0.0167, (2.393) | 0.0993, (1.648) | 0.0836, (1.730) | 0.0208, (2.311) | 0.0333, (2.128) | 0.2406, (1.173) |
| X = 20 | Y = 3 | 0.0018, (3.125) | 0.0835, (1.731) | 0.0104, (2.563) | 0.0089, (2.616) | 0.0807, (1.747) | 0.0464, (1.992) |
| X = 30 | Y = 3 | 0.0006, (3.411) | 0.0257, (2.230) | 0.0094, (2.598) | 0.0005, (3.498) | 0.0333, (2.128) | 0.0055, (2.775) |
| X = 40 | Y = 3 | 0.0004, (3.510) | 0.0257, (2.231) | 0.0060, (2.747) | 0.0012, (3.244) | 0.0779, (1.763) | 0.0057, (2.765) |
| X = 50 | Y = 3 | 0.0007, (3.394) | 0.0257, (2.230) | 0.0083, (2.638) | 0.0001, (3.811) | 0.0216, (2.298) | 0.0024, (3.039) |
| X = 60 | Y = 3 | 0.0002, (3.690) | 0.0573, (1.901) | 0.0022, (3.062) | 0.0001, (4.013) | 0.0085, (2.632) | 0.0025, (3.021) |

The shaded area highlights all comparisons that do **not** lead to a significant effect at the 10-percent level.

TABLE 7

Pairwise comparison of transfers to active Players B in the helping game per return transfer by a given investment

| Wilcoxon signed-rank tests: p-values, z in parentheses | | | | |
|--|----------------------------|------------------------------------|-----------------|---------------------|
| <i>Investment (X)</i> | <i>Return Transfer (Y)</i> | <i>Baseline & Anticipation</i> | <i>Baseline</i> | <i>Anticipation</i> |
| X = 10 vs. X = 20 | Y = 0 | 0.5638, (0.577) | 0.3173, (1.000) | 0.9738, (0.033) |
| X = 10 vs. X = 30 | Y = 0 | 1.0000, (0.000) | 1.0000, (0.000) | 1.0000, (0.000) |
| X = 30 vs. X = 40 | Y = 0 | 0.5729, (0.564) | 0.3173, (1.000) | 0.9738, (0.033) |
| X = 40 vs. X = 50 | Y = 0 | 0.5458, (0.604) | 0.3173, (1.000) | 0.9738, (0.033) |
| X = 50 vs. X = 60 | Y = 0 | 0.5547, (0.591) | 0.3173, (1.000) | 0.9738, (0.033) |
| X = 10 vs. X = 20 | Y = 1 | 0.9300, (0.088) | 0.3300, (0.974) | 0.3697, (0.897) |
| X = 10 vs. X = 30 | Y = 1 | 1.0000, (0.000) | 1.0000, (0.000) | 1.0000, (0.000) |
| X = 30 vs. X = 40 | Y = 1 | 0.3353, (0.963) | 0.5275, (0.632) | 0.4944, (0.683) |
| X = 40 vs. X = 50 | Y = 1 | 0.5547, (0.591) | 1.0000, (0.000) | 0.5455, (0.605) |
| X = 50 vs. X = 60 | Y = 1 | 0.7264, (0.350) | 0.3173, (1.000) | 0.9834, (0.021) |
| X = 10 vs. X = 20 | Y = 2 | 0.2889, (1.061) | 0.6717, (0.424) | 0.1027, (1.632) |
| X = 10 vs. X = 30 | Y = 2 | 1.0000, (0.000) | 1.0000, (0.000) | 1.0000, (0.000) |
| X = 30 vs. X = 40 | Y = 2 | 0.0608, (1.875) | 0.5924, (0.535) | 0.0478, (1.9799) |
| X = 40 vs. X = 50 | Y = 2 | 0.1452, (1.457) | 0.3547, (0.925) | 0.2273, (1.207) |
| X = 50 vs. X = 60 | Y = 2 | 0.2854, (1.068) | 0.2838, (1.072) | 0.6218, (0.493) |
| X = 10 vs. X = 20 | Y = 3 | 0.2689, (1.106) | 0.5455, (0.605) | 0.1048, (1.622) |
| X = 10 vs. X = 30 | Y = 3 | 1.0000, (0.000) | 1.0000, (0.000) | 1.0000, (0.000) |
| X = 30 vs. X = 40 | Y = 3 | 0.1545, (1.424) | 0.3064, (1.023) | 0.0179, (2.367) |
| X = 40 vs. X = 50 | Y = 3 | 0.4364, (0.778) | 0.3547, (0.925) | 0.7970, (0.257) |
| X = 50 vs. X = 60 | Y = 3 | 0.8501, (0.189) | 0.3805, (0.877) | 0.6069, (0.514) |

The shaded area highlights all comparisons that do **not** lead to a significant effect at the 10-percent level.

TABLE 8

Pairwise comparison of transfers to active Players B in the helping game per return transfer and per investment by treatment

| Mann-Whitney rank-sum tests: p-values, z in parentheses | | |
|---|----------------------------|----------------------------------|
| <i>Investment (X)</i> | <i>Return Transfer (Y)</i> | <i>Baseline vs. Anticipation</i> |
| X = 10 | Y = 0 | 0.7333, (0.341) |
| X = 20 | Y = 0 | 0.3346, (0.965) |
| X = 30 | Y = 0 | 0.9812, (0.024) |
| X = 40 | Y = 0 | 0.3113, (1.012) |
| X = 50 | Y = 0 | 0.6242, (0.490) |
| X = 60 | Y = 0 | 0.3116, (1.012) |
| X = 10 | Y = 1 | 0.5311, (0.626) |
| X = 20 | Y = 1 | 0.3373, (0.960) |
| X = 30 | Y = 1 | 0.3213, (0.992) |
| X = 40 | Y = 1 | 0.2961, (1.045) |
| X = 50 | Y = 1 | 0.4058, (0.831) |
| X = 60 | Y = 1 | 0.5307, (0.627) |
| X = 10 | Y = 2 | 0.6731, (0.422) |
| X = 20 | Y = 2 | 0.3165, (1.002) |
| X = 30 | Y = 2 | 0.5060, (0.665) |
| X = 40 | Y = 2 | 0.4055, (0.832) |
| X = 50 | Y = 2 | 0.3594, (0.916) |
| X = 60 | Y = 2 | 0.4473, (0.760) |
| X = 10 | Y = 3 | 0.7100, (0.372) |
| X = 20 | Y = 3 | 0.3164, (1.002) |
| X = 30 | Y = 3 | 0.3699, (0.897) |
| X = 40 | Y = 3 | 0.2288, (1.203) |
| X = 50 | Y = 3 | 0.2430, (1.167) |
| X = 60 | Y = 3 | 0.2332, (1.192) |

The shaded area highlights all comparisons that do **not** lead to a significant effect at the 10-percent level.

TABLE 9

Pairwise comparison of relative return transfers by active Players B in the trust game by a given increase in the investment by 10

| Wilcoxon signed-rank tests: p-values, z in parentheses | | | |
|--|------------------------------------|-----------------|---------------------|
| <i>Investment (X)</i> | <i>Baseline & Anticipation</i> | <i>Baseline</i> | <i>Anticipation</i> |
| X = 10 vs. X = 20 | 0.0839, (1.728) | 0.0294, (2.178) | 1.0000, (0.000) |
| X = 10 vs. X = 30 | 0.3173, (1.000) | 0.1573, (1.414) | 1.0000, (0.000) |
| X = 30 vs. X = 40 | 0.5637, (0.577) | 0.3173, (1.000) | 0.1573, (1.414) |
| X = 40 vs. X = 50 | 0.1005, (1.642) | 1.0000, (0.000) | 0.0987, (1.651) |
| X = 50 vs. X = 60 | 0.5073, (0.663) | 0.4644, (0.732) | 1.0000, (0.000) |

The shaded area highlights all comparisons that do **not** lead to a significant effect at the 10-percent level.

TABLE 10

Pairwise comparison of relative return transfers for a given investment by active Players B in the trust game by treatment

| Mann-Whitney rank-sum tests: p-values, z in parentheses | |
|---|----------------------------------|
| <i>Investment (X)</i> | <i>Baseline vs. Anticipation</i> |
| X = 10 | 0.0022, (3.056) |
| X = 20 | 0.0107, (2.551) |
| X = 30 | 0.0279, (2.199) |
| X = 40 | 0.0927, (1.681) |
| X = 50 | 0.4370, (0.605) |
| X = 60 | 0.6622, (0.437) |

The shaded area highlights all comparisons that do **not** lead to a significant effect at the 10-percent level.

TABLE 11

Possible payoffs for Players B including trust game given the played strategy of helpers in the Anticipation Treatment

| <i>Investment (X)</i> | <i>Return Transfer (Y)</i> | <i>Possible Payoff in Trust Game for Player B</i> | <i>Mean Help Decision (standard errors in parentheses)</i> | <i>Mean Help received (helpers' decision tripled)</i> | <i>Total possible Payoff for Player B</i> |
|---|----------------------------|---|--|---|---|
| 10 | 0 | 130 | 1.27, (4.4584) | 3.82 | 133.82 |
| 10 | 1 | 129 | 3.73, (5.4404) | 11.18 | 140.18 |
| 10 | 2 | 128 | 11.16, (16.259) | 33.48 | 161.48 |
| 10 | 3 | 127 | 14.73, (18.775) | 44.18 | 171.18 |
| 20 | 0 | 160 | 1.61, (7.7556) | 4.84 | 164.84 |
| 20 | 1 | 159 | 4.64, (9.8408) | 13.91 | 172.91 |
| 20 | 2 | 158 | 13.45, (19.069) | 40.36 | 198.36 |
| 20 | 3 | 157 | 16.68, (21.851) | 50.05 | 207.05 |
| 30 | 0 | 190 | 0.52, (2.1184) | 1.57 | 191.57 |
| 30 | 1 | 189 | 5.14, (9.6727) | 15.41 | 204.41 |
| 30 | 2 | 188 | 13.45, (19.906) | 40.36 | 228.36 |
| 30 | 3 | 187 | 19.05, (25.356) | 57.14 | 244.14 |
| 40 | 0 | 220 | 0.70, (2.5479) | 2.11 | 222.11 |
| 40 | 1 | 219 | 5.93, (10.500) | 17.80 | 236.80 |
| 40 | 2 | 218 | 16.18, (24.779) | 48.55 | 266.55 |
| 40 | 3 | 217 | 20.84, (28.356) | 62.52 | 279.52 |
| 50 | 0 | 250 | 1.27, (4.9526) | 3.82 | 253.82 |
| 50 | 1 | 249 | 6.39, (12.289) | 19.16 | 268.16 |
| 50 | 2 | 248 | 16.41, (24.831) | 49.23 | 297.23 |
| 50 | 3 | 247 | 20.39, (27.252) | 61.16 | 308.16 |
| 60 | 0 | 280 | 1.61, (6.7967) | 4.84 | 284.84 |
| 60 | 1 | 279 | 6.39, (13.113) | 19.16 | 298.16 |
| 60 | 2 | 278 | 16.34, (24.240) | 49.02 | 327.02 |
| 60 | 3 | 277 | 19.95, (25.231) | 59.86 | 336.86 |
| The orange areas highlight the maximal profit a player B could have earned given an investment. | | | | | |

TABLE 12*Possible payoffs for investors in the trust game given the played strategy of trustees*

| <i>Investment (X)</i> | <i>Baseline</i> | | | <i>Anticipation</i> | | |
|--|---|---|---|---|---|---|
| | <i>Mean Trustees Relative Transfer (Y) (standard errors in parentheses)</i> | <i>Mean Trustees Absolut Return (Y*X)</i> | <i>Total possible Payoff for investor</i> | <i>Mean Trustees Relative Transfer (Y) (standard errors in parentheses)</i> | <i>Mean Trustees Absolut Return (Y*X)</i> | <i>Total possible Payoff for investor</i> |
| 0 | ---- | ---- | 100.00 | ---- | ---- | 100.00 |
| 10 | 0.4091, (0.8541) | 4.09 | 94.09 | 1.27, (1.0320) | 12.72 | 102.72 |
| 20 | 0.6818, (0.6463) | 13.63 | 93.63 | 1.27, (0.7673) | 25.45 | 105.45 |
| 30 | 0.7727, (0.8125) | 23.18 | 93.18 | 1.27, (0.7673) | 38.18 | 108.18 |
| 40 | 0.8182, (0.8528) | 32.72 | 92.72 | 1.18, (0.7327) | 47.27 | 107.27 |
| 50 | 0.8182, (0.8528) | 40.90 | 90.90 | 0.95, (0.8439) | 47.72 | 97.72 |
| 60 | 0.8636, (0.9902) | 51.81 | 91.81 | 0.95, (0.8985) | 57.27 | 97.27 |
| The orange areas highlight the maximal profit an investor could have earned in the trust game in the two conditions. | | | | | | |