

Bruni, Luigino; Tufano, Fabio

**Working Paper**

## The value of vulnerability: The transformative capacity of risky trust

CeDEx Discussion Paper Series, No. 2015-19

**Provided in Cooperation with:**

The University of Nottingham, Centre for Decision Research and Experimental Economics (CeDEx)

*Suggested Citation:* Bruni, Luigino; Tufano, Fabio (2015) : The value of vulnerability: The transformative capacity of risky trust, CeDEx Discussion Paper Series, No. 2015-19, The University of Nottingham, Centre for Decision Research and Experimental Economics (CeDEx), Nottingham

This Version is available at:

<https://hdl.handle.net/10419/129831>

**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.*

Discussion Paper No. 2015-19

Luigino Bruni and  
Fabio Tufano

November 2015

**The value of vulnerability**  
**The transformative**  
**capacity of risky trust**

CeDEx Discussion Paper Series  
ISSN 1749 - 3293



CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

The Centre for Decision Research and Experimental Economics was founded in 2000, and is based in the School of Economics at the University of Nottingham.

The focus for the Centre is research into individual and strategic decision-making using a combination of theoretical and experimental methods. On the theory side, members of the Centre investigate individual choice under uncertainty, cooperative and non-cooperative game theory, as well as theories of psychology, bounded rationality and evolutionary game theory. Members of the Centre have applied experimental methods in the fields of public economics, individual choice under risk and uncertainty, strategic interaction, and the performance of auctions, markets and other economic institutions. Much of the Centre's research involves collaborative projects with researchers from other departments in the UK and overseas.

Please visit <http://www.nottingham.ac.uk/cedex> for more information about the Centre or contact

Suzanne Robey  
Centre for Decision Research and Experimental Economics  
School of Economics  
University of Nottingham  
University Park  
Nottingham  
NG7 2RD  
Tel: +44 (0)115 95 14763  
Fax: +44 (0) 115 95 14159  
[suzanne.robey@nottingham.ac.uk](mailto:suzanne.robey@nottingham.ac.uk)

The full list of CeDEX Discussion Papers is available at

<http://www.nottingham.ac.uk/cedex/publications/discussion-papers/index.aspx>

# The value of vulnerability

## The transformative capacity of risky trust

LUIGINO BRUNI AND FABIO TUFANO\*

**Abstract:** *The ‘grammar of trust’ is one the most explored loci in game theory and behavioural economics. However, still much needs to be understood about the nature of trust in non-enforceable, personalised interactions, in markets and within organizations. This experimental study aims at contributing to the understanding of new dimensions of trust by exploring how risky and vulnerable trust may foster a trustee’s behavioural change. It investigates trustee’s positive reciprocity when the intentional vulnerability of the trustor is both manifestly salient and dependent upon the trustee’s revealed type. The results support the transformative nature of risky trust by showing the capacity of vulnerable trust to generate more fairness and reciprocity in untrustworthy people.*

**JEL classification:** C72; C92; D01

**Keywords:** experiment; gift-exchange game; organization; trust; vulnerability

### 1. Introduction

Trust is the key resource for both civil society and market economy. The Latin *fides* (trust) meant rope, to express the essential social role of trust: linking people one another, the bond of society (Locke (1954[1660], 213). When trust is missing or weakened, market economy is undermined and the whole fabric of civil society becomes insecure and fragile. For this reason, at

---

\*Bruni: Dipartimento di Scienze economiche, politiche e delle lingue moderne, LUMSA, Via Pompeo Magno 22, 00192 Roma, Italy (e-mail: l.bruni@lumsa.it). Tufano: CeDEx, School of Economics, University of Nottingham, University Park, Nottingham NG7 2RD, UK (e-mail: fabio.tufano@nottingham.ac.uk). Part of the research was done when Bruni and Tufano were at the University of Milano-Bicocca (Italy). The authors thank Irene Brundia for help and suggestions at the early stage of the design, and for lab assistance. We thank also Luca Stanca for his inputs and support especially in the initial phases of the study. Research funding from the Center for Interdisciplinary Studies in Economics, Psychology and Social Sciences (University of Milan-Bicocca) is gratefully acknowledged.

the very beginning of modern Political Economy, trust was a key issue in classical philosophers and economists such as Adam Smith, David Hume, or Antonio Genovesi (Bruni and Sugden, 2000). After more than two centuries of almost total eclipses, in recent years trust is again among the interest of economists (e.g., Glaeser et al., 2000; Alesina and La Ferrara, 2002; Butler et al., forthcoming).

Despite behavioural economics and social sciences are more and more dispelling the ‘grammar of trust’ (e.g., Fehr, 2009), still much needs to be understood about the nature of trust in non-enforceable, personalised interactions in markets and within organizations.

The paper aims at contributing to such understanding, by exploring the reciprocity capacity of *vulnerable* trust from an experimental organizational viewpoint (e.g., Camerer and Weber, 2012). It investigates trustee’s positive reciprocity when the intentional vulnerability of the trustor is both manifestly *salient* and clearly *dependent* upon the trustee’s revealed type.

That genuine (i.e., not purely self-interested) trust involves vulnerability is acknowledged in interdisciplinary literature on trust (e.g., Rousseau et al., 1998). Vulnerability, however, is often interpreted in a negative way, as involuntary ‘exposure’ to other people’s action or events, normally due to the lack of resources, rights, capabilities, empowerment or freedom.<sup>1</sup> The human development is usually measured in terms of reduction or elimination of this negative and unchosen vulnerability. At the same time, some philosophers and social scholars claim also for a positive value of vulnerability, related to the inherent fragility associated to good life (i.e. reciprocal love, friendship, relational goods, and civic commitment in general): “Mutual activity, feeling, and awareness are such a deep part of what love and friendship *are*... that there is anything worthy of the name of love or friendship left, when the shared activities and the forms of communication that express it are taken away” (Nussbaum, 1986, p. 344). However, “these components of the good life are going to be minimally self-sufficient. And they will be vulnerable in an especially deep and dangerous way” (1986, *Ib.*). Without this sort of positive vulnerability, human life does not flourish fully while economic activities and organizations do not fulfil entirely their potential (Bruni, 2012).

By resorting to an experimental gift-exchange game, our study explores a specific dimension of this positive (although risky and often costly) vulnerability in interpersonal relations: that is, its capacity of eliciting positive reciprocity when the trustor’s vulnerability depends upon the trustee’s past behaviour and consequently revealed type. To the best of our knowledge, no experiment has been conducted so far to investigate the effects of trusting untrustworthy individuals, when this specific risk is common knowledge and is made salient. In fact, we study experimentally the transformative capacity of risky trust: that is, the possibility that trusting untrustworthy individuals

---

<sup>1</sup> On the concept of ‘exposure’ see also Pelligra (2007).

may change their responses from non-reciprocal, unfair, responses to fair ones. More to the point, our research question is as it follows: does vulnerable trust has a transformative positive effect on untrustworthy individuals? Our laboratory evidence says unambiguously *yes*. Salient vulnerability of trust has sizeable effects on trustee's transfers. Both fairness and reciprocity of trustee's behaviour increase when vulnerability is made salient. Overall, in our experimental setting, there is consistent evidence that vulnerability of trust leads to transformative effects on the behaviour of untrustworthy trustees, who tend to be more reciprocating.

The paper is structured as follows. Section 2 presents the relevant literature and the theoretical framework of our analysis. Section 3 describes the experimental design. Section 4 discusses the predictions to be tested. Section 5 presents the results. Section 6 concludes with a discussion of the main findings and implications of the analysis.

## **2. The Risk of Vulnerability**

A large number of lab and field experiments show that reciprocal behaviour in economic interactions is influenced by fairness concerns and by others' intentions and motivations. Fairness is by far the most explored issues in the already abundant social-preference literature (e.g., Fehr and Gächter, 2000; Sobel, 2005; and Fehr and Schmidt, 2006). Models differ, however, with respect to *how* fairness is defined and then measured. Two possible classes of models can be identified: the first class entails models focusing on distributional concerns or inequity aversion preferences (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000), while the second one counts models centred on the role of intentions, which may be perceived as either kind or unkind (e.g., Rabin, 1993; Dufwenberg and Kirchsteiger, 2004). This second class of models links together intention-fairness and reciprocity, whenever an agent's utility functions is sensitive of perceived (un)kindness of her opponent. In these latter models, actions with identical outcomes may elicit different reciprocating responses depending on how agents' intentions are interpreted by their opponents. Some models explore a role of intentions on the basis of the 'attribution hypothesis:' that is, by comparing intentional choices with non-intentional ones, and interpreting as (un)kind only intentional actions (e.g., McCabe, 2003; Falk et al., 2008). Other models (e.g., Andreoni et al., 2002; Falk et al., 2003) use as proxy of (un)kindness the strategy space of the agent: if, for their opponent's benefit, the agent chooses a dominated course of action when a more rational was available, this choice signals kindness (and *vice versa*). Variant of these intention-based models are Stanca et al. (2009), where kind intentions are associated with non-strategic behaviour, and Bauernshuster et al. (2013), where

kindness depends on frames (i.e., more competitive frames reduce perceived kindness). Bolton and Ockenfels (2005) show that in their experiment both (distributional and intentional) theories account for agents' evaluation of fairness.

Our experimental study relates to intention-based models. To investigate trust, it analyses individual behaviour in a two-player symmetric gift-exchange game (Stanca et al., 2009). In such a game, both players are endowed with  $E$  monetary units. The first mover can either keep the initial endowment for her or send ' $x \leq E$ ' monetary units from her initial endowment to the second mover. Then,  $x$  is multiplied by a factor  $m$  by the experimenter. Thus, the second mover receives ' $x \cdot m$ ' monetary units; he can then either keep his initial endowment  $E$  or send ' $y \leq E$ ' from his initial endowment to the first mover. Thus, the payoffs for the first and second movers are equal to ' $E - x + m \cdot y$ ' and ' $E + m \cdot x - y$ ', respectively. In line with the literature on trust (e.g., Bohnet, 2008), we will refer to the amount sent forward by the first mover as a proxy of trust, while the amount sent back by the second mover as a proxy of trustworthiness. (Accordingly, in what follows, we use the terms 'trustor' and 'trustee' as interchangeable with 'first mover' and 'second mover,' respectively). The relevant dimension of the first mover's action to be interpreted as "kind" by the second mover is her intentional, costly, and full informed choice of becoming *vulnerable* with respect to the second mover's action (i.e., being exposed to the risk of exploitation by the second mover).

Berg et al. (1995), Pillutla et al. (2003), Malhotra (2004) and Strassmair (2009) deal with issues similar to our own.<sup>2</sup> These studies do not find any significant effect of the trustor's exposure on the trustee's degree of reciprocity. These results depend on critical features of the experimental designs, which – we believe – were not purposely built to disentangle the transformative effect of vulnerability on trustee's reciprocity. In particular, in Berg et al. (1995) and Pillutla et al. (2003), whenever trustors took higher risks by sending larger portions of their endowment, they provided greater benefit to trustees as a consequence. Thus, there is no possibility to know whether trustees reciprocated because they valued the risks trustors had undertaken, or for distributional reasons - or for both.

Malhotra's (2004) study is the closest to our own study, with a specific acknowledgement of the role of vulnerability. He finds no significant impact of the trustor's risk on the trustee's trustworthiness, which is instead significantly affected by the benefit provided to them by the trustor. In contrast with previous studies, Malhotra (2004) maintained separate the effect of the trustor's risk from the trustee's benefit by treating these factors as control variables of the decision-making context, but still the trustor's vulnerability is not manifestly salient and is not linked to the trustee's behaviour and revealed type. In fact, in his study the only dimension of risk of exposure is

---

<sup>2</sup> Cialdini (1993) and Reagan (1971) deal with risk and reciprocity, but the issue of vulnerability is not among their analyses.

measured by the variation in the trustor's material payoff: exactly this variation is used in every interaction to determine the intentionality and vulnerability of trust. We claim that this strategy is not able to disentangle the specific risk of the trustor's vulnerability inherent to their interaction with a given (untrustworthy) trustee.

Strassmair (2009) studies the effect of expected future rewards on actions' perceived kindness and relative reciprocal responses. To this aim, she varies the probability for the trustee to make a return transfer. In Strassmair's (2009) low treatment the probability for the trustee of deciding on their return transfer is 50%; in her high treatment the respective probability is 80%. Therefore, in the low treatment the second mover would perceive the first mover as kinder than in the high treatment, *ceteris paribus*, and therefore return more in the former than in the later treatment whenever asked to make a decision. The results, however, did not show this correlation. In fact, in Strassmair's (2009) experiment participants seem more sensitive to distributional fairness rather than their opponent's expectation of future rewards. This therefore suggests that second movers are insensitive to the vulnerability of first movers' trust.

The vulnerability of trust is explained as a disposition of the trustor to accept the risk to be betrayed by the trustee (Baier, 1986). This represents the fundamental element that distinguishes genuine or authentic trust from other forms of general trust as reliance; the authentic trust, unlike the non-genuine, emerges only in context of human relationships in which the presence of people – rather than machines (i.e., computer) – explains the possibility to feel betrayed rather than the mere possibility to be disappointed. Thus, we consider such a presence as a first necessary condition for the vulnerability of trust as we intend it. However, it does not constitute a *sufficient* condition to explain it. As also Holton (2004) underlines, one person could choose to undertake actions based on trust without taking the risk to be betrayed, that is without undertaking any vulnerability. This leads to what we consider a second necessary condition for the vulnerability of trust: that is, the risk of trusting depends on the trustee's type and their level of trustworthiness.

Upon maintaining the first condition in each and every experimental treatment as in previous related studies (e.g., Blount 2005, Falk et al. 2008, Stanca et al. 2009; Stanca, 2010), we design our experiment to explore the effect of the second condition whenever made manifestly salient. In fact, we posit that an element that unequivocally supports the trustor's willingness to be potentially betrayed by the trustee is the risk to trust an individual who, by making an unfair decision, proved to be untrustworthy in a recent interaction with a third person. We hypothesise that this kind of trustor's risk, that we call *vulnerability*, may have a transformative effect on the reciprocal response of an untrustworthy trustee. This is what we investigate by means of our experiment as detailed below.



### 3. Experimental Design and Procedures

In the experiment we resort to a two-player symmetric gift-exchange game (Stanca et al., 2009). This game involves two stages. In each stage, only one player is active. Both players – say Ann and Bob, for the sake of convenience – receive an initial endowment of 5 tokens each. In Stage 1, the active player Ann decides how many tokens of her initial endowment to keep for her and how many tokens (only integers could be disposed) to send to her co-player Bob. Then, the ‘ $x$ ’ tokens sent by Ann are multiplied by 3 by the experimenter. Therefore Bob receives ‘ $3 \cdot x$ ’. In Stage 2, Bob decides how many tokens of his initial endowment to keep for him and how many tokens (only integers could be disposed) to send to Ann. Then, the ‘ $y$ ’ tokens sent by Bob are multiplied by 3 by the experimenter. Therefore, Ann receives ‘ $3 \cdot y$ ’. In summary, Ann’s final payoff of the game is ‘ $5 - x + 3 \cdot y$ ’, while Bob’s final payoff is ‘ $5 + 3 \cdot x - y$ ’. The final game payoff can vary between a minimum of 0 to a maximum of 20 tokens.

Within each experimental session, the game was played three times. We refer to those three times as Game 1, 2 and 3, respectively. Players learned about the games step by step. Players’ roles were fixed across games: that is, if Ann (*resp.* Bob) was the first (*resp.* second) mover by being active in Stage 1 (*resp.* Stage 2) of Game 1, she (*resp.* he) would have been the first (*resp.* second) mover in Game 2 and 3, too. Game 1-3 were each played with a different co-player. In other words, a stranger matching protocol was in place. In Stage 2 of Game 3 we applied a variant of the strategy method similar to the one implemented in Fischbacher et al. (2001). The active player in such a stage had to make a *set of decisions*: that is, one for each of the 6 possible numbers of tokens they could receive from their co-player. Thus, a schedule of tokens to be sent back was elicited (i.e., a vector of tokens to be sent back). In all other stages, the decision method was always applied entailing a *single decision* about how many tokens from the initial endowment a player sent to their co-player. Accordingly, with the only exception of Game 3, the active player in Stage 2 was always made aware of the tokens sent by their co-player when taking their decision.

Players’ choices were rewarded via a lottery incentive mechanism. One between Game 1, 2 and 3 was randomly selected and played for real. If either Game 1 or 2 was selected, the two single decisions were used to determine the final game payoffs for the two co-players. If Game 3 was selected, then the schedule of the player active in Stage 2 became relevant and the decision corresponding to the actual number of tokens received was played for real. The monetary earnings from the experiment were obtained by converting the final payoff of the randomly selected game in euros by applying an exchange rate of 2 tokens equal to 1 euro, plus 5 euros as show-up fee.

There are two treatments in the experiment: the Information treatment (I-treatment *for short*) and No-Information treatment (N-treatment *for short*). The experimental manipulation across treatments entails the disclosure or not of information about the second mover's choice in Game 1. In fact, in Game 2 and 3 of the I-treatment, first movers are informed whether their co-player made either a "fair" or an "unfair" choice in Game 1, while the second movers were made aware that their co-players received such information. By contrast, in Game 2 and 3 of the N-treatment, no information about second movers' choices in Game 1 was disclosed. A second mover's choice in Game 1 was labelled as "fair" (*resp.* "unfair") if they sent to their co-player a number of token larger than or equal to (*resp.* lower than) the tokens they received. For the sake of the exposition, we would refer to players making a fair (*resp.* unfair) choice as trustworthy (*resp.* untrustworthy). The two treatments were identical in all other respects.

Experimental participants were invited by e-mail using a list of voluntary potential candidates. Upon arrival, participants were randomly assigned to a computer workstation. The experiment started with instructions read aloud by the experimenters to set ground rules. Then, participants were led step by step by computerized instructions. First, Game-1 instructions were displayed and subjects learned about their role (i.e., either first or second mover), which stayed the same throughout the experiment. Then, participants made their choices and moved on Game 2. After going through Game 3, subjects learned the randomly selected game to be played for real, their choice, their opponent's choice and their earnings. The experiment ended with a standard background questionnaire.

Two-hundred eight students (of whom 59.62 percent enrolled in undergraduate degrees) drawn from a range of academic disciplines (with Business and Economics summing up respectively to 50.96 percent and 17.79 percent of the whole sample) participated in our experiment, which took place at the Experimental Economics Laboratory of the University of Milano-Bicocca (Italy). Our experiment lasted one hour on average and was computerized in z-Tree (Fischbacher, 2007). Participants were paid individually and anonymously at the end of each experimental session.<sup>3</sup>

---

<sup>3</sup> All participants received the total sum of the actual earnings from the experiment as described in the main text plus a € 5.00 show-up fee. Total payments ranged between € 5.00 and € 15.00 with an average payment equal to € 9.40 (standard deviation of € 3.12).

## 4. Predictions

The present study focuses on the behavioural implications of the vulnerability of trust. To explore those implications, it is necessary to concentrate the attention on the second mover. Therefore, in what follows, the predictions are stated exactly with regard to second movers' behaviour.

Upon assuming that players are purely self-interested and this is common knowledge, the second mover who is at the game terminal node will always send zero tokens to the first mover. By backward induction, the first mover will rationally choose to not send any token to the second mover. Therefore, the standard equilibrium prediction is that both players will send zero tokens. However, if standard social preferences are postulated, second movers may optimally choose a non-zero transfer leading to possible predictions with positive transfers of tokens. Those predictions should hold irrespectively of our treatment manipulation.

Differently, if second mover's preferences show concerns for the vulnerability of trust, the amount of tokens sent back by them should be higher when the first mover's vulnerability is salient.

*Hypothesis 1.* If second movers' preferences present concerns for the vulnerability of trust, their behavioural strategy implies higher numbers of tokens transferred in the I-treatment.

Whether or not second movers' strategies imply higher numbers of tokens transferred when vulnerability is salient, it remains to be investigated if such an implication regards *tout court* both trustworthy and untrustworthy players.

*Hypothesis 1.a (resp. 1.b).* Assuming concerns for the vulnerability of trust, the behavioural strategy of a trustworthy (*resp.* untrustworthy) second movers implies higher numbers of tokens transferred in the I-treatment.

In the I-treatment, higher transfers of tokens by second movers do not imply per se a higher share of fair choices and, consequently, of trustworthy players. In other words, when vulnerability is salient, it is conceivable that second movers' behavioural strategies could imply more generous but not yet fair transfers, which would leave unchanged the share of trustworthy players.

*Hypothesis 2.* If second movers' preferences show concerns for the vulnerability of trust, the share of trustworthy second movers is larger in the I-treatment.

Whatever is the impact of trust vulnerability on the implied level of transfers and on the share of trustworthy people, its transformative capacity may also impact on the reciprocity attitudes of second movers. In fact, both an increase in the transfer levels and a higher share of trustworthy people may result simply from an upward shift of second movers' behavioural strategies. By contrast, a change in the second movers' reciprocity attitudes would require a different correlation between first and second movers' transfers, or in other words, a change in the 'slope' of second movers' behavioural strategies.

*Hypothesis 3.* Assuming concerns for the vulnerability of trust, there is higher correlation between first and second movers' transfers in the I-treatment.

Whether this is the case at the (second-mover) population level, or not, it is still important to explore separately the strategic patterns of the trustworthy (*resp.* untrustworthy) players.

*Hypothesis 3a (resp. 3b).* If second mover's preferences present concerns for the vulnerability of trust, the correlation between a trustworthy (*resp.* untrustworthy) player's transfer and their received tokens is higher in the I-treatment.

## 5. Results

Table 1 reports summary statistics by treatment, game and type of players. First of all, both first and second movers transfer on average non-zero amounts of tokens to their co-player. Therefore, as it should be expected on the basis of the abundant experiment literature, a prediction made upon assuming purely self-interested players should be rejected. In Game 1 of the N-treatment (*resp.* I-treatment), first movers sent on average 3.077 (*resp.* 2.846) tokens to their respective co-players; second movers responded by sending back on average 1.885 (*resp.* 1.712), which are still positive but lower than what full reciprocity would imply. A set of Wilcoxon rank-sum tests ( $p$ -values  $> 0.40$ ) demonstrate that there is no statistically significant difference in Game 1 between the average amounts sent by first (*resp.* second) movers across treatments, showing a successful random assignment of participants in treatments and roles.

Game 2 of the N-treatment was a close replica of the previous game outcomes. On average, first movers transferred 3.192 tokens to their co-players who responded by sending back 1.885

tokens. In the I-treatment a slight change in the average behaviour was reported in Game 2: first and second movers transferred 2.365 and 1.981 tokens, respectively.

Game 3 presents a different overall picture. First movers transferred on average 2.135 (*resp.* 2.731) tokens to second movers who in turn sent back 1.474 (*resp.* 1.962) tokens in the N-treatment (*resp.* I-treatment).

TABLE 1—SUMMARY STATISTICS OF TOKEN TRANSFERS BY TREATMENT, GAME AND TYPE OF PLAYER

	N-treatment						No. Obs.
	Game 1		Game 2		Game 3		
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	
First Mover's transfer	3.077	1.702	3.192	1.783	2.135	1.961	52
Second Mover's transfer	1.885	1.843	1.885	1.916	1.474	0.996	52

	I-treatment						No. Obs.
	Game 1		Game 2		Game 3		
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	
First Mover's transfer	2.846	1.564	2.365	1.794	2.731	2.097	52
Second Mover's transfer	1.712	1.73	1.981	1.873	1.962	1.108	52

*Notes:* In Game 3, the raw data for calculating the mean and standard deviation for the second mover's transfer were obtained by averaging the individual transfers elicited by the strategy method.

Given its central role in addressing our research questions, Game 3 data requires a more detailed exploration. Figure 1 shows the distribution of transfers as resulting from the strategy method elicitation procedure. By comparing the left panel (N-treatment data) with the right panel (I-treatment data) of Figure 1, it seems apparent a rightward shift of the medians of token transferred (see the thick black lines in the grey boxes) when moving from the left to the right panel; moreover, the distributions of the tokens sent back seem less dispersed as demonstrated by the reduction of the grey-shaded areas.

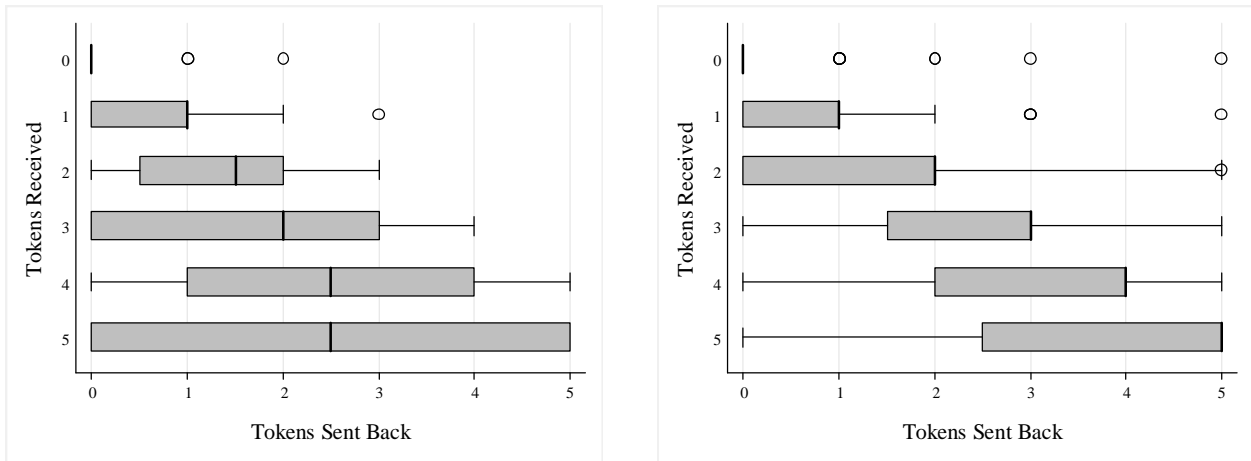


FIGURE 1. TOKENS SENT BACK IN THE N-TREATMENT (*left panel*) AND I-TREATMENT (*right panel*) FROM GAME 3 (STRATEGY METHOD)

Notes: The left (*resp.* right) hinge of each grey box reports the 25th (*resp.* 75th) percentile; the thick black line represents the median; the vertical line of the left (*resp.* right) whiskers shows lower (*resp.* higher) adjacent values; the spherical markers show outside values.

It is now time to turn to our hypothesis testing. By comparing across treatments second movers' transfers in Game 3, it results a significant increase in the amount of token sent in the I-treatment (Wilcoxon rank-sum tests:  $z = -2.275$ ;  $p = 0.0229$ ). To further test the robustness of this evidence, Table 2 reports our econometric analysis. The OLS estimation of Model 1 demonstrates that participants in the I-treatment make higher token transfers in Game 3 than participants in the N-treatment; moreover, trustworthy second-mover participants do transfer more tokens. The estimation of Model 1 does not show any history-dependent pattern: in fact, the transfers of players in Game 3 do not depend on the transfers received in previous games. The OLS estimation of Model 2 confirms the insights from Model 1 while a possible significant interaction between a treatment effect and trustworthiness is not corroborated. The ordered logit estimation of Model 3 provides further robustness to the evidence already obtained by examining Model 1 and 2. Overall, there is a significant increase in the amount of token sent in the I-treatment even when controlling for participants' trustworthiness, the history of plays and possible interaction effects.

TABLE 2—ECONOMETRIC ANALYSIS OF TOKEN TRANSFERS IN GAME 3 (STRATEGY METHOD)

Model	1		2		3	
Estimation Procedure	OLS		OLS		Ordered Logit	
Dependent variable	Average transfer		Average transfer		Individual transfer	
Clustering	No		No		Yes	
I-treatment (dummy)	0.568***	(0.207)	0.571**	(0.284)	0.548**	(0.215)
First mover's transfer:						
• Game 1	0.091	(0.064)	0.091	(0.064)	0.108	(0.076)
• Game 2	0.048	(0.057)	0.048	(0.057)	0.054	(0.065)
Trustworthy (dummy)	0.532**	(0.208)	0.535*	(0.290)	0.597***	(0.217)
I-treatment × Trustworthy (dummy)			-0.006	(0.408)		
Observations	104		104		624	
Log likelihood					-1017.989	
R-squared	0.127		0.127			

Notes: In Game 3, the raw data for calculating the dependent variables respectively in Model 1 and 2 were obtained by averaging second-mover individual transfers elicited by the strategy method.

Therefore, when the vulnerability of risk is made salient by providing additional information to players, second movers do increase their transfers. This inference is confirmed by Figure 2, which shows both effect sizes and absolute treatment differences across the whole schedule of choices.

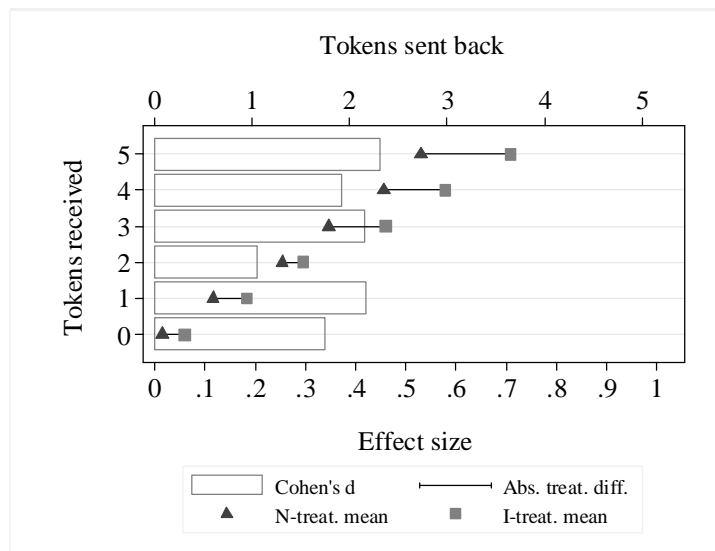


FIGURE 2. TREATMENT EFFECTS IN SECOND MOVERS' TRANSFERS (GAME 3, STRATEGY METHOD)

Notes: The Cohen's *d* can be read on the bottom horizontal axis labeled 'Effect size,' while the absolute treatment difference in the mean token transfers between the N-treatment and the I-treatment can be read on the top horizontal axis labeled "Tokens sent back." The vertical axis report the tokens received by second movers.

*Result 1.* The null hypothesis of no treatment differences can be significantly rejected in favour of sizable effects of the vulnerability of trust on second movers' transfers.

Overall, second movers' transfers significantly increase when vulnerability of trust is salient. However, it is reasonable to inquire if that is the case both for trustworthy and untrustworthy second movers. By comparing the strategy method average transfers of trustworthy second movers, a significant difference results across treatments: that is, in the I-treatment trustworthy second movers send back more tokens on average than in the N-treatment (Wilcoxon rank-sum test:  $z = 2.072$ ;  $p = 0.038$ ). A similar comparison of average transfers of untrustworthy second movers shows a similar qualitative pattern, which turns out marginally non-significant (Wilcoxon rank-sum test:  $z = 1.570$ ;  $p = 0.117$ ).

*Result 1a and 1b.* When vulnerability of risk is salient, trustworthy second movers transfer significantly higher amounts of tokens. An increase in token transfers, which just miss to meet the significance threshold, is detected for untrustworthy second movers.

In order to assess the transformative capacity of the vulnerability of trust, we compare the number of (un)trustworthy second movers across treatments. In Game 1, due to random assignment, there was no significant difference in second movers' trustworthiness (Wilcoxon rank-sum test:  $z = 0.392$ ;  $p = 0.695$ ). In Game 3, after calculating the average fairness of each stated strategy (which ranges between 0 and 1), a Wilcoxon rank-sum test ( $z = 2.252$ ;  $p = 0.024$ ) comparing average strategy fairness across treatment shows a significant higher trustworthiness in the I-treatment.

*Result 2.* By comparing average strategy trustworthiness, the I-treatment presents a significant difference. Thus, vulnerable trust has a transformative capacity.

Let now turn to the test of hypothesis 3. Figure 3 reports the distribution of Spearman correlation coefficients in Game 3 between first and second movers' transfers in the I-treatment (left panel) and the N-treatment (right panel), respectively. By inspecting Figure 3, higher Spearman correlation coefficients appear only slightly more frequently in the I-treatment than in the N-treatment. These graphical insights are corroborated by statistical testing. In fact, by statistically comparing the Spearman correlations among all second movers in Game 3 across treatments no significant difference can be detected (Wilcoxon rank-sum test:  $z = 1.024$ ;  $p = 0.3056$ ).



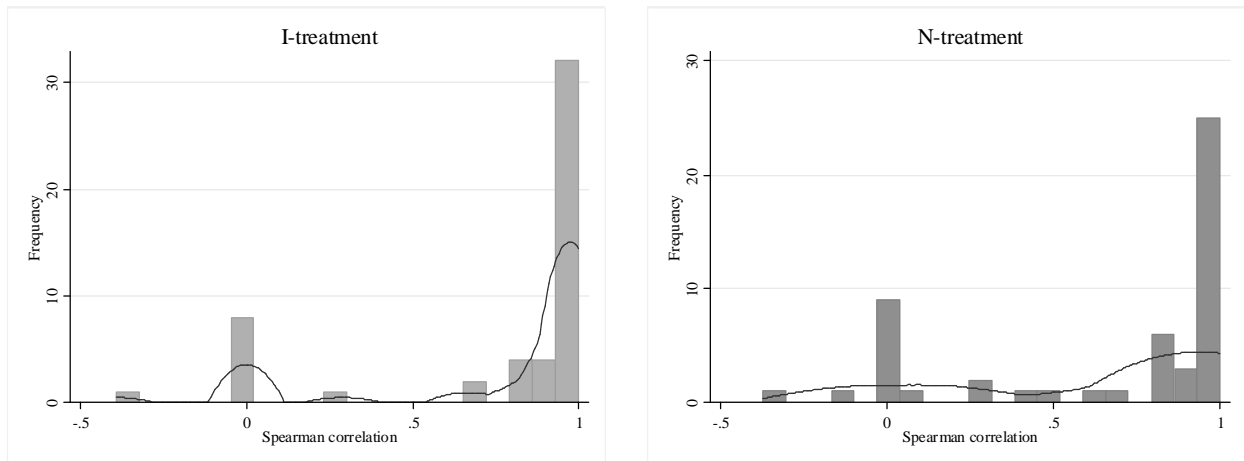


FIGURE 3. SPEARMAN CORRELATIONS ACROSS TREATMENTS (GAME 3, STRATEGY METHOD)

*Notes:* In Game 3, the Spearman coefficients are calculated by considering the correlation between the hypothetical first-mover token transfer and the second-mover one as stated on the second mover's transfer schedule.

Notwithstanding, it could be possible that even if at the aggregate level there is no difference across treatments in the correlations between first and second movers' transfers, there might be treatment differences in the strategic patterns of trustworthy and untrustworthy players, respectively. To investigate such a possibility, it is necessary to disaggregate the data. Figure 4 presents the distribution of spearman correlation coefficients between first and second movers' transfers across treatments and types (i.e., trustworthy and untrustworthy), respectively. By focusing on the top panels of Figure 4 that displays transfer correlations of pairs with trustworthy second movers, it seems that there is no sharp difference in the correlation coefficients and, if anything, the N-treatment shows slightly higher coefficients. However, the graphical evidence changes quite dramatically when looking at the bottom panels of Figure 4. In fact, the correlation coefficients for the I-treatment with untrustworthy second movers are higher than the respective ones for the N-treatment. The statistical analysis confirms the absence of any significant difference across treatments when considering pairs with a trustworthy second mover (Wilcoxon rank-sum test:  $z = 0.107$ ;  $p = 0.9145$ ). A similar comparison of correlation coefficients across treatments when considering pairs with an untrustworthy second mover confirms the qualitative insights from the graphical analysis: that is, the spearman correlation coefficient are higher for untrustworthy second movers in the I-treatment, but the difference across treatments turns out marginally non-significant (Wilcoxon rank-sum test:  $z = 1.584$ ;  $p = 0.1215$ ).

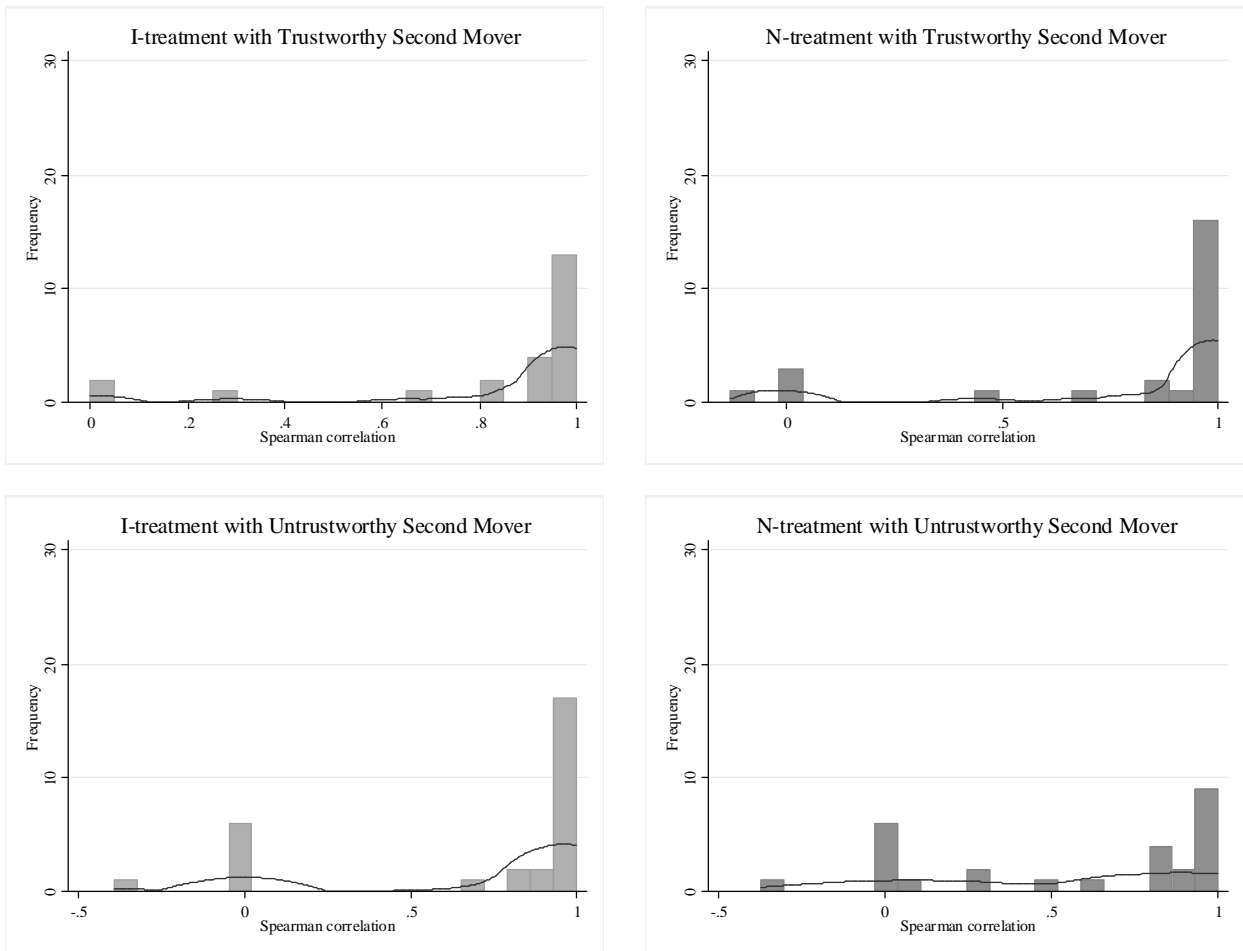


FIGURE 4. SPEARMAN CORRELATIONS ACROSS TREATMENTS AND SECOND MOVERS' TYPES (GAME 3, STRATEGY METHOD)

Notes: In Game 3, the Spearman coefficients are calculated by considering the correlation between the hypothetical first-mover token transfer and the second-mover one as stated on the second mover's transfer schedule.

*Result 3.* Overall, there is no evidence that the correlation between first and second movers' transfers are higher in the I-treatment. At more disaggregate level, the correlation coefficients between tokens received and transfers made by trustworthy second movers cannot be statistically distinguished across treatments; however, the correlation between tokens received and untrustworthy second movers' transfers shows a treatment difference, which just miss to meet the significance threshold.

Therefore, the evidence collected is strongly suggestive that the vulnerability of trust transforms the reciprocity tendency of untrustworthy second movers, who increase the ratio between their transfer and the tokens received.

## 6. Concluding remarks

Our experiment has consistently shown across treatments that when the vulnerability of the trustors (or, first movers) is linked to the trustee's type and is made salient by providing relevant information to the players, the trustees (or, second movers) do change their behaviour by increasing the amount of tokens transferred. In those circumstances, both trustworthy and untrustworthy trustees make more generous transfers, and the degrees of fairness and reciprocity of trustees' behaviour rise. In synthesis, the transformative nature of vulnerable trust finds consistent support as shown by its capacity of generating more positive, fair and reciprocal transfers by trustees.

Both linking the trustor's vulnerability to the trustee's type and its salience have been the key elements in the study. Accordingly, we postulate that the lack of these elements of trust vulnerability in previous experimental studies was the main reason of the absence of any significant role of the trustor's vulnerability in explaining the trustee's behaviour.

It is easy to envision relevant fields where our results may suggest policy implications and, in general, reflections and suggestions.

One domain is management. Our result may be relevant in the so-called "managerial subsidiarity", according to which the manager has to intervene in the decisions of a team only for those activities that would be worse without her *subsidiary* intervention (Melé, 2004). But for subsidiarity management to function it is essential that workers and work groups feel genuine trust, and therefore vulnerable. To make subsidiarity effective, it is important that the management should really trust the work group, and it should not want to control or 'contractualise' the entire process to prevent abuse of trust. If, however, those who are given "delegation" perceive that in fact the trust given to them is only instrumental to profit maximization, subsidiarity can stop producing its effects in eliciting creativity and innovation. Then, a key issue in subsidiarity management is the resilience after a crisis due to the abuse of the trustee, when untrustworthiness is known and the organization wants to keep its culture of trust. Our results support the effectiveness of subsidiarity and offer strength to the value of giving new trust to workers that have shown themselves as untrustworthy.

Subsidiarity is essential also in education, where teachers have to create an environment of genuine trust in order to elicit responsibility and freedom. Trust games are the common settings in most educational programs – in schools and in programs for adults with problematic and untrustworthy past experiences. To trust children, young and people with past experiences of untrustworthiness is a key issue on which mostly depends the success of the education process. In particular, our results suggest that making salient the vulnerable risk of the trustor (i.e., teacher or social worker) may produce a "transformative" effect on the trustee (Horsburgh, 1960). Vulnerable

and risky trust can have a therapeutic value that can cure relational failures.

Finally we hope that our study will stimulate replications and further research in order to accumulate systematic knowledge on the grammar of trust, especially in non-enforceable, personalised interactions.

## References

- Alesina, Alberto and Eliana La Ferrara. 2002. Who Trusts Others? *Journal of Public Economics*, 85(2): 207-34.
- Andreoni, J., Brown, P., and Vesterlund, L. 2002. What makes an allocation fair? Some experimental evidence. *Games and Economic Behavior* 40, 1–24.
- Baier, A. 1986. Trust and Antritrust. *Ethics*. 96: 231-260.
- Bauernschuster, S., Falck, O., and Große, N. 2013. When trustors compete for the favour of a trustee—A laboratory experiment. *Journal of Economic Psychology* 34, 133-147.
- Berg, J., Dickhaut, J., and McCabe, K. 1995. Trust, reciprocity, and social history. *Games and Economic Behavior* 10, 122–142.
- Bohnet, Iris. 2008. Trust in experiments. *The New Palgrave Dictionary of Economics*. Second Edition. Eds. Steven N. Durlauf and Lawrence E. Blume. Palgrave Macmillan, 2008. *The New Palgrave Dictionary of Economics Online*. Palgrave Macmillan.
- Blount, S. 1995. When social outcomes aren't fair: the effect of causal attributions on preferences. *Organizational Behavior and Human Decision Processes* 63, 131–144.
- Bolton, G.E., and Ockenfels, A. 2000. ERC: a theory of equity, reciprocity and competition. *American Economic Review* 90, 166–193.
- Bolton, G. E., and Ockenfels, A. 2005. A stress test of fairness measures in models of social utility. *Economic Theory* 25(4), 957-982.
- Bruni L. and R. Sugden. 2000. Moral canals. Trust and social capital in the work of Hume, Smith and Genovesi. *Economics and Philosophy* 16, 21-45.
- Bruni, L. 2012. *The wound and the blessing. Economics, interpersonal relations, happiness*, Newcity, New York.
- Butler, J., P. Giuliano and L. Guiso. Fothcoming. The Right Amount of Trust. *Journal of the European Economic Association*.
- Camerer, C.F., and R.A. Weber. (2012). Experimental organizational economics. In Robert Gibbons and John Roberts (Eds.), *Handbook of Organizational Economics*. Princeton University Press.

- Cialdini, R.B. *Influence: Science and Practice*, 5th ed. Pearson. Boston.
- Dufwenberg, M., and Kirchsteiger, G. 2004. A theory of sequential reciprocity. *Games and Economic Behavior* 47, 268–298.
- Falk, A., Fehr, E., and Fischbacher, U. 2003. On the nature of fair behaviour. *Economic Inquiry* 41, 20–26.
- Falk, A., Fehr, E., and Fischbacher, U., 2008. Testing theories of fairness—intentions matter. *Games and Economic Behavior* 62, 287–303
- Fehr, Ernst. 2009. On the Economics and Biology of Trust. *Journal of the European Economic Association* 7(2-3): 235-266.
- Fehr, E., and Gächter, S. 2000. Fairness and retaliation: the economics of reciprocity. *Journal of Economic Perspectives* 14 (3), 159–181.
- Fehr, E., & Schmidt, K. M. 1999. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114, 817-868.
- Fehr, E., Schmidt, K.M., 2006. The economics of fairness, reciprocity and altruism: experimental evidence. In: Kolm, S.C., Ythier, J.M. (Eds.), *Handbook of the Economics of Giving, Altruism and Reciprocity*, vol. 1. Elsevier, Amsterdam, pp. 615–691.
- Glaeser, E., D. Laibson, J.A. Scheinkman and C.L. Soutter. 2000. Measuring Trust. *Quarterly Journal of Economics* 115(3): 811-846
- Holton, R. 1994. Deciding to Trust, Coming to Believe. *Australian Journal of Philosophy* 72. 63-76.
- Horsburgh, H.J.N. 1960. The Ethics of Trust. *The Philosophical Quarterly*.10. 343-354.
- Locke, J. 1954[1660]. *Essays of the law of nature*. Clarendon, Oxford.
- McCabe, K.A., Rigdon, M.L., and Smith, V.L. 2003. Positive reciprocity and intentions in trust games. *Journal of Economic Behavior and Organization* 52, 267–275.
- Malhotra, D. 2004. Trust and reciprocity decision: The differing perspectives of trustors and trusted parties. *Organizational Behavior and Human Decision Processes* 94. 61-73.
- Melé, D. 2004. Exploring the principle of subsidiarity in organizational forms. *Journal of Business Ethics*, 60, 293-305.
- Nussbaum, M. 1986. *The fragility of goodness*. CUP. Cambridge.
- Pelligra, V. 2007. The Not-so-Fragile Fragility of Goodness: The Responsive Quality of Fiduciary Relationships. In P.L. Porta and L. Bruni, *Handbook of Happiness in Economics*. Cheltenham. Edward Elgar.
- Pillutla, M., Malhotra, D., and Murnighan, J. K. 2003. Attributions of trust and the calculus of reciprocity. *Journal of Experimental Social Psychology* 39, 448–455.

- Potter, N.N., 2002. How can I be trusted? A virtue theory of trustworthiness. Rowman & Littlefield Publishers, Inc. Boston.
- Rabin, M. 1993. Incorporating fairness into game theory and economics. *American Economic Review* 83, 1281–1302.
- Regan, R. T. (1971). Effects of a favor and liking on compliance. *Journal of Experimental Social Psychology*. 7. 627–639.
- Rousseau, D.M., Sitkin, S.B., Burt, R.S., Camerer, C. 1998. Not so different after all: a cross-discipline view of trust. *Academy of Management Review* 23, 393-404.
- Sobel, J. 2005. Interdependent preferences and reciprocity. *Journal of Economic Literature* 43, 392–436.
- Stanca, L., Bruni, L., Corazzini, L. 2009. Testing theories of reciprocity: Do motivations matter? *Journal of Economic Behavior & Organization*. 71. 233-245.
- Stanca, L. 2010. How to be kind? Outcomes versus intentions as determinant of fairness. *Economic Letters*. 106. 19-21.
- Strassmair, C. 2009. Can intention spoil the kindness of a gift? An experimental study. Discussion paper 2009/1. Department of Economics, University of Munich.

# The Value of Vulnerability

## The Transformative Capacity of Risky Trust

LUIGINO BRUNI AND FABIO TUFANO

### Experimental Instructions

#### I. Oral Instructions

##### ORIGINAL IN ITALIAN

Benvenuto e grazie per la partecipazione a questo esperimento.

Durante l'esperimento non è consentito parlare o comunicare in alcun modo con gli altri soggetti. Se in qualsiasi momento hai una domanda alza la mano e uno degli assistenti verrà a risponderti.

Seguendo attentamente le istruzioni potrai guadagnare un ammontare in gettoni che dipenderà dalla tue scelte e da quelle degli altri soggetti.

L'esperimento è diviso in fasi. Ogni fase è composta da un gioco.

Al termine dell'esperimento, i gettoni che avrai guadagnato saranno convertiti in Euro al tasso di cambio di 2 gettoni = 1 Euro. Solo un gioco tra tutti quelli cui avrai partecipato ti verrà remunerato ed esso sarà scelto in modo casuale alla fine dell'esperimento.

##### ENGLISH TRANSLATION

Welcome and thank you for participating in this experiment.

During the experiment, you are not allowed to talk, or communicate in any other way with other subjects. If you have questions, please, raise your hand and an assistant will address your question.

By following carefully the instructions you may earn an amount of tokens that will depend upon your choices and those of the other subjects.

The experiment consists of several phases. Each phase involves a game.

At the end of the experiment, the tokens you will have earned will be converted in euros with an exchange rate equal to 2 tokens = € 1.00. Only one game among the ones you played will be paid for real and it will be randomly selected.

Ad ogni gioco partecipano 2 soggetti, tu e l'altro giocatore.

All'inizio dell'esperimento a te e all'altro giocatore, in modo casuale, saranno assegnati il ruolo di giocatore A o di giocatore B. Questi ruoli rimarranno fissi durante tutto l'esperimento sebbene, all'inizio di ogni gioco, verranno formulate nuove coppie di soggetti. Es.: se tu sei un giocatore A, rimarrai con questo ruolo per tutto l'esperimento e, ad ogni gioco, interagirai con un diverso giocatore B.

Ciascun soggetto, pertanto, in tutti i giochi, interagirà con soggetti diversi senza mai conoscerne l'identità.

For each game, 2 subjects will take part in: that is, you and another player.

At the beginning of the experiment, you and the other player will be randomly assigned to the role of player A and B, respectively. These roles will stay the same throughout the experiment even though, at the start of each game, new pairs of players will be formed. For instance, if you are a player 'A,' you will play in such a role throughout the experiment and, in each game, you will interact with a different player 'B.'

Thus, each subject will interact in every game with different subjects without knowing their identity.



## II. Computerized Instructions

The screenshots (I-treatment) taken from the original z-Tree code used in our experimental sessions are reproduced below together with the English translation.

### a. Player A

SCREENSHOT 1A



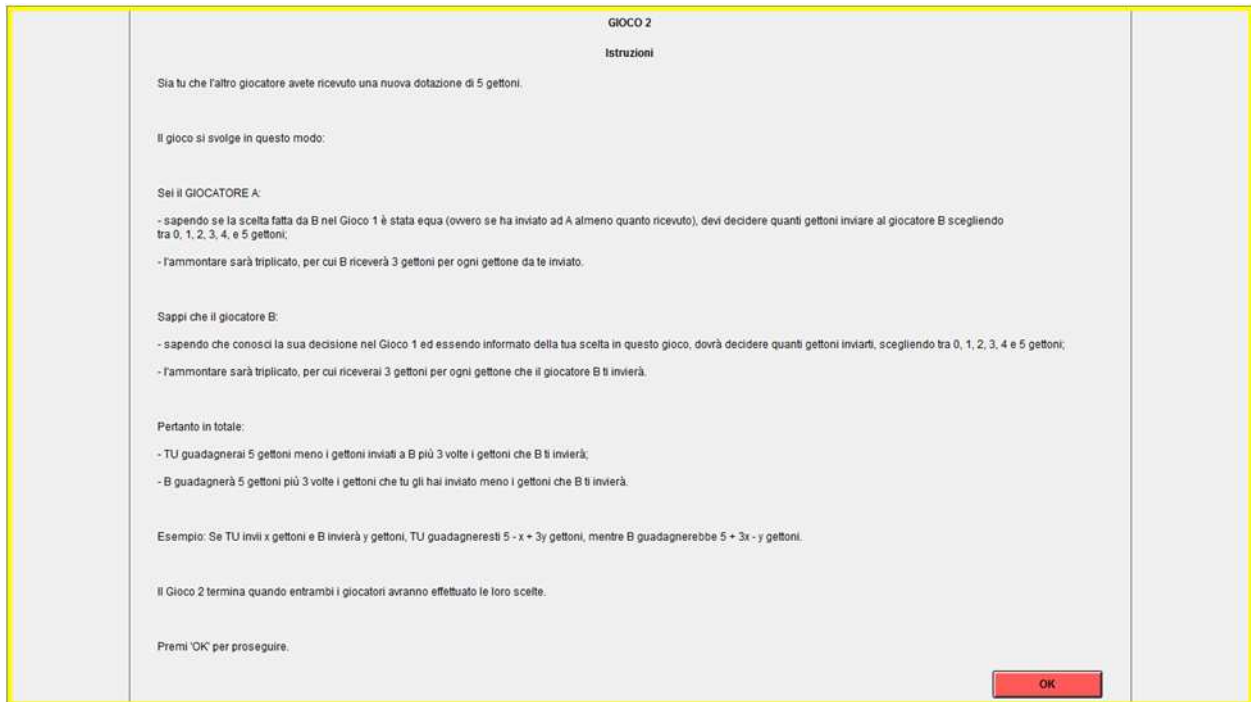
*English translation (top-down):* Game 1 || Instructions || Both you and the other player have received an endowment of 5 tokens || The game goes as follows: || You are PLAYER A || - you should decide how many tokens to send to player B by choosing between 0, 1, 2, 3, 4, and 5 tokens; || - the amount will be tripled, thus B will receive 3 tokens per each token you sent. || Notice that player B: || - informed about your choice, will decide how many tokens to send you by choosing between 0, 1, 2, 3, 4, and 5 tokens; || - the amount will be tripled, thus you will receive 3 tokens per each token B will send you. || Therefore, in total: || YOU will earn 5 tokens minus the tokens sent to B plus 3 times the tokens B will send you || B will earn 5 tokens plus 3 times the tokens you sent them minus the tokens B will send you || For example: if YOU send  $x$  tokens and B will send you  $y$  tokens, YOU would earn  $5 - x + 3y$  tokens, while B would earn  $5 + 3x - y$  tokens. || Game 1 ends when both players would have made their choices. || Press 'OK' to continue. || OK

SCREENSHOT 2A



English translation (top-down): Game 1 || How many tokens do you want to send to PLAYER B? || Confirm

### SCREENSHOT 3A



*English translation (top-down):* Game 2 || Instructions || Both you and the other player have received an endowment of 5 tokens || The game goes as follows: || You are PLAYER A || - knowing if B's choice in Game 1 was fair (that is, if you send to A at least as much as you received), you should decide how many tokens to send to player B by choosing between 0, 1, 2, 3, 4, and 5 tokens; || - the amount will be tripled, thus B will receive 3 tokens per each token you sent. || Notice that player B: || - knowing that you were aware of their decision in Game 1 and being informed of your choice in this game, will have to decide how many tokens to send you by choosing between 0, 1, 2, 3, 4, and 5 tokens; || - the amount will be tripled, thus you will receive 3 tokens per each token B will send you. || Therefore, in total: || YOU will earn 5 tokens minus the tokens sent to B plus 3 times the tokens B will send you || B will earn 5 tokens plus 3 times the tokens you sent them minus the tokens B will send you || For example: if YOU send  $x$  tokens and B will send you  $y$  tokens, YOU would earn  $5-x+3y$  tokens, while B would earn  $5+3x-y$  tokens. || Game 2 ends when both players would have made their choices. || Press 'OK' to continue. || OK

SCREENSHOT 4A



*English translation (top-down):* Game 2 || Notice that in Game 1 PLAYER B was unfair. || How many tokens do you want to send to PLAYER B? || Confirm

*Notes:* If player B in Game 1 sent to A at least as much as they received, the word “unfair” in the above text would have been substituted by “fair.”

## SCREENSHOT 5A



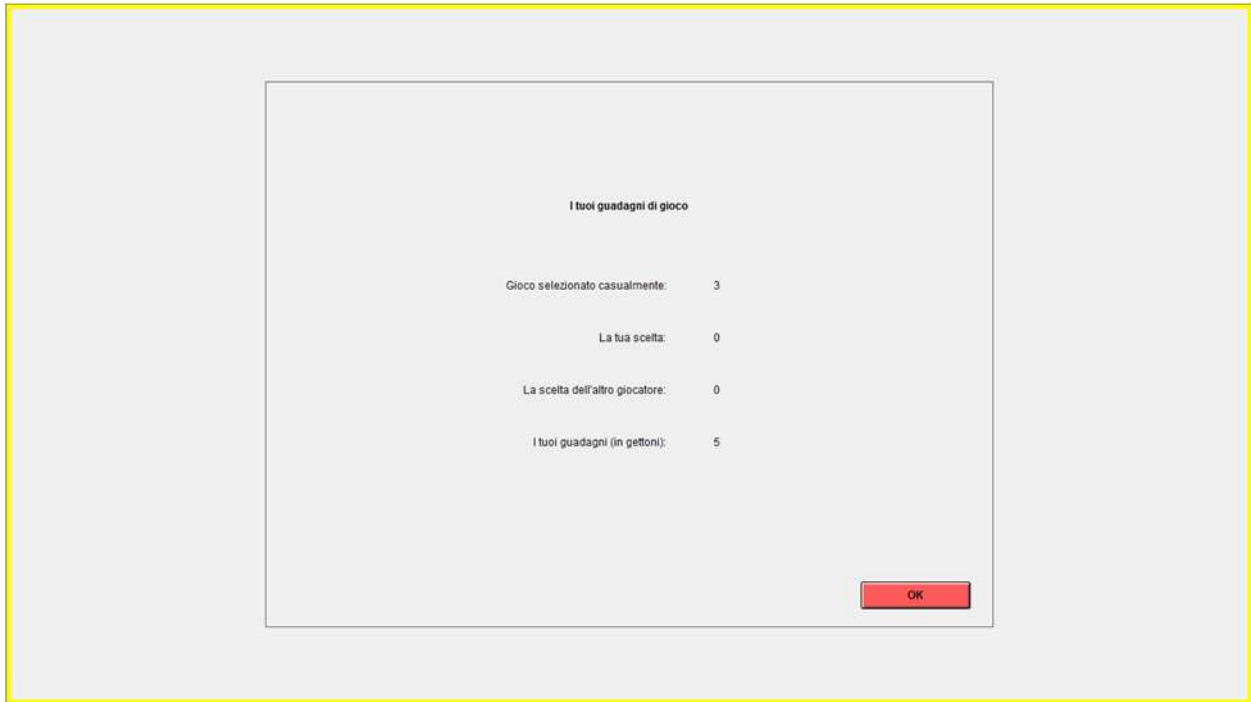
*English translation (top-down):* Game 3 || Instructions || Both you and the other player have received an endowment of 5 tokens || The game goes as follows: || You are PLAYER A || - knowing if B's choice in Game 1 was fair (that is, if you send to A at least as much as you received), you should decide how many tokens to send to player B by choosing between 0, 1, 2, 3, 4, and 5 tokens; || - the amount will be tripled, thus B will receive 3 tokens per each token you sent. || Notice that player B: || - knowing that you were aware of their decision in Game 1 and being informed of your choice in this game, will have to decide how many tokens to send you by choosing between 0, 1, 2, 3, 4, and 5 tokens; || - the amount will be tripled, thus you will receive 3 tokens per each token B will send you. || Therefore, in total: || YOU will earn 5 tokens minus the tokens sent to B plus 3 times the tokens B will send you || B will earn 5 tokens plus 3 times the tokens you sent them minus the tokens B will send you || For example: if YOU send  $x$  tokens and B will send you  $y$  tokens, YOU would earn  $5-x+3y$  tokens, while B would earn  $5+3x-y$  tokens. || Game 3 ends when both players would have made their choices. || Press 'OK' to continue. || OK

SCREENSHOT 6A



English translation (top-down): Game 3 || Notice that in Game 1 PLAYER B was unfair. || How many tokens do you want to send to PLAYER B? || Confirm

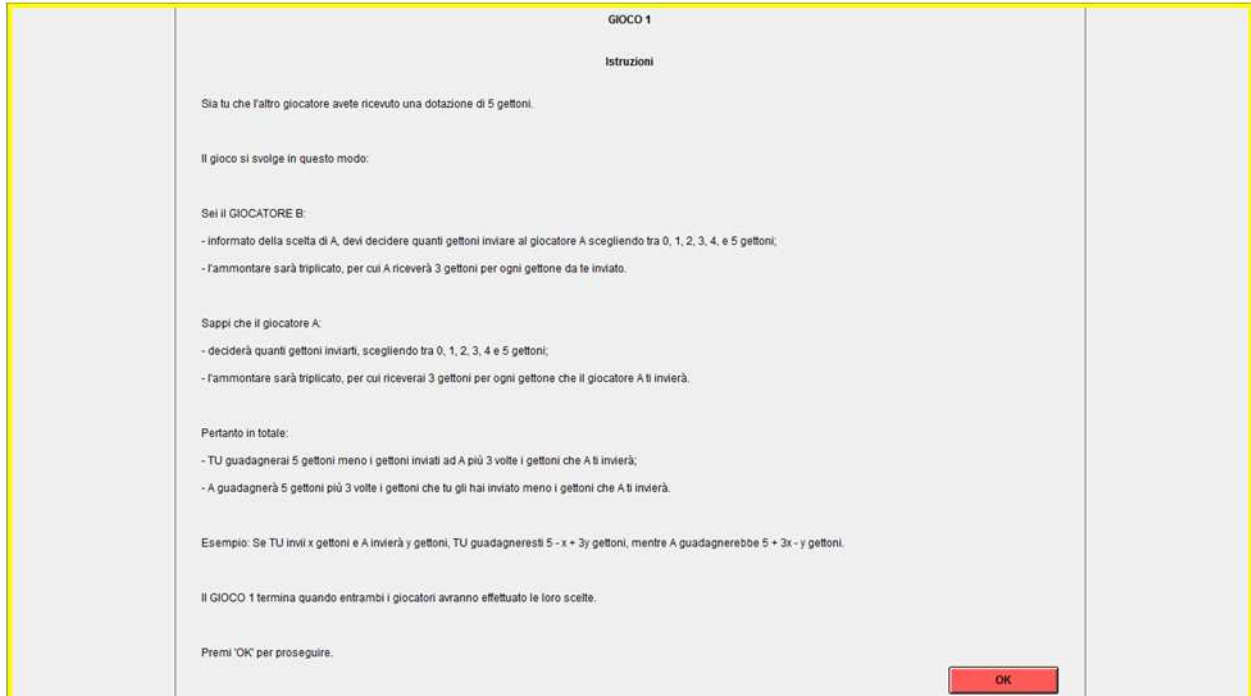
SCREENSHOT 7A



*English translation (top-down):* Your game earnings || Game randomly selected:3 || Your choice: 0 || The choice of the other player: 0 || Your earnings (in tokens): 5 || OK

## b. Player B

SCREENSHOT 1B



*English translation (top-down):* Game 1 || Instructions || Both you and the other player have received an endowment of 5 tokens || The game goes as follows: || You are **PLAYER B** || - once you learn about A's choice, you will be asked to decide how many tokens to send to player A by choosing between 0, 1, 2, 3, 4, and 5 tokens; || - the amount will be tripled, thus A will receive 3 tokens per each token you sent. || Notice that player A: || - will decide how many tokens to send you by choosing between 0, 1, 2, 3, 4, and 5 tokens; || - the amount will be tripled, thus you will receive 3 tokens per each token A sent you. || Therefore, in total: || YOU will earn 5 tokens minus the tokens sent to A plus 3 times the tokens A will send you || A will earn 5 tokens plus 3 times the tokens you sent them minus the tokens A will send you || For example: if YOU send x tokens and A will send you y tokens, YOU would earn  $5 - x + 3y$  tokens, while A would earn  $5 + 3x - y$  tokens. || Game 1 ends when both players would have made their choices. || Press 'OK' to continue. || OK

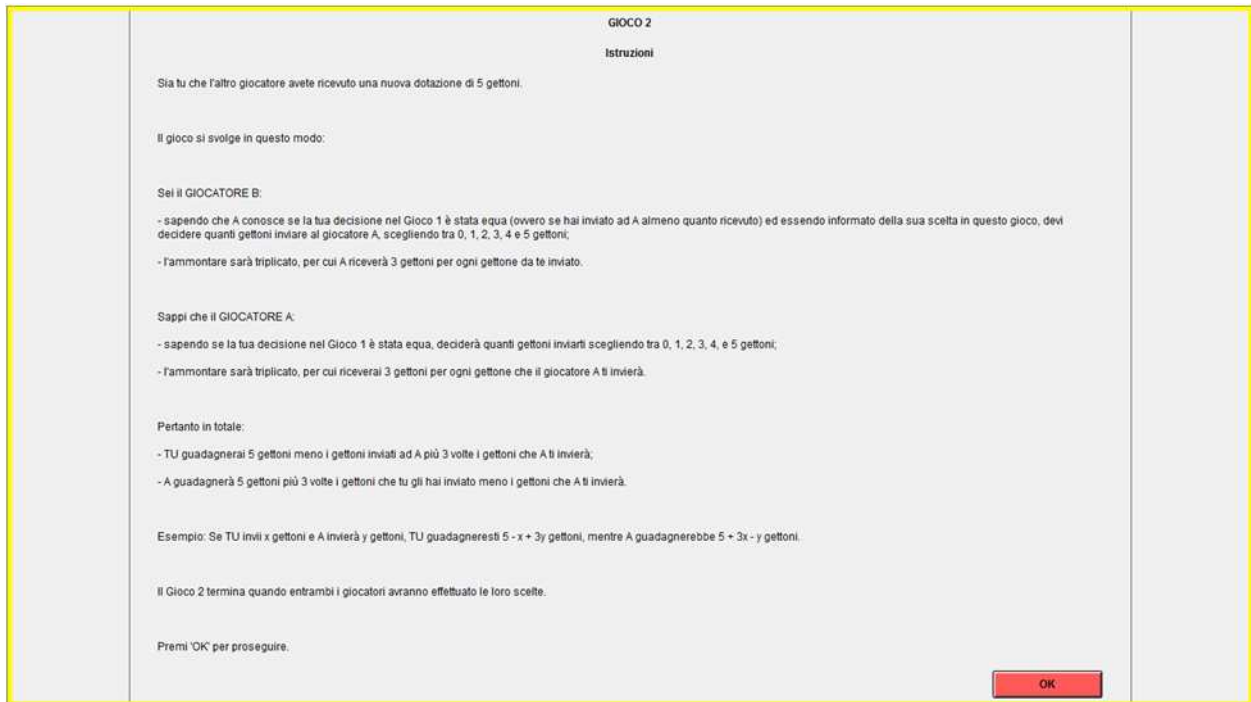


SCREENSHOT 2B



English translation (top-down): Game 1 || In this game, PLAYER A sent you a number of tokens equal to 5 || How many tokens do you want to send to PLAYER A? || Confirm

### SCREENSHOT 3B



*English translation (top-down):* Game 2 || Instructions || Both you and the other player have received an endowment of 5 tokens || The game goes as follows: || You are PLAYER B || - knowing both that A is aware if your decision in Game 1 was fair (that is, if you send to A at least as much as you received) and their choice in this game, you should decide how many tokens to send to player A by choosing between 0, 1, 2, 3, 4, and 5 tokens; || - the amount will be tripled, thus A will receive 3 tokens per each token you sent. || Notice that player A: || - knowing if your decision in Game 1 was fair, will decide how many tokens to send you by choosing between 0, 1, 2, 3, 4, and 5 tokens; || - the amount will be tripled, thus you will receive 3 tokens per each token A sent you. || Therefore, in total: || YOU will earn 5 tokens minus the tokens sent to A plus 3 times the tokens A will send you || A will earn 5 tokens plus 3 times the tokens you sent them minus the tokens A will send you || For example: if YOU send  $x$  tokens and A will send you  $y$  tokens, YOU would earn  $5 - x + 3y$  tokens, while A would earn  $5 + 3x - y$  tokens. || Game 2 ends when both players would have made their choices. || Press 'OK' to continue. || OK

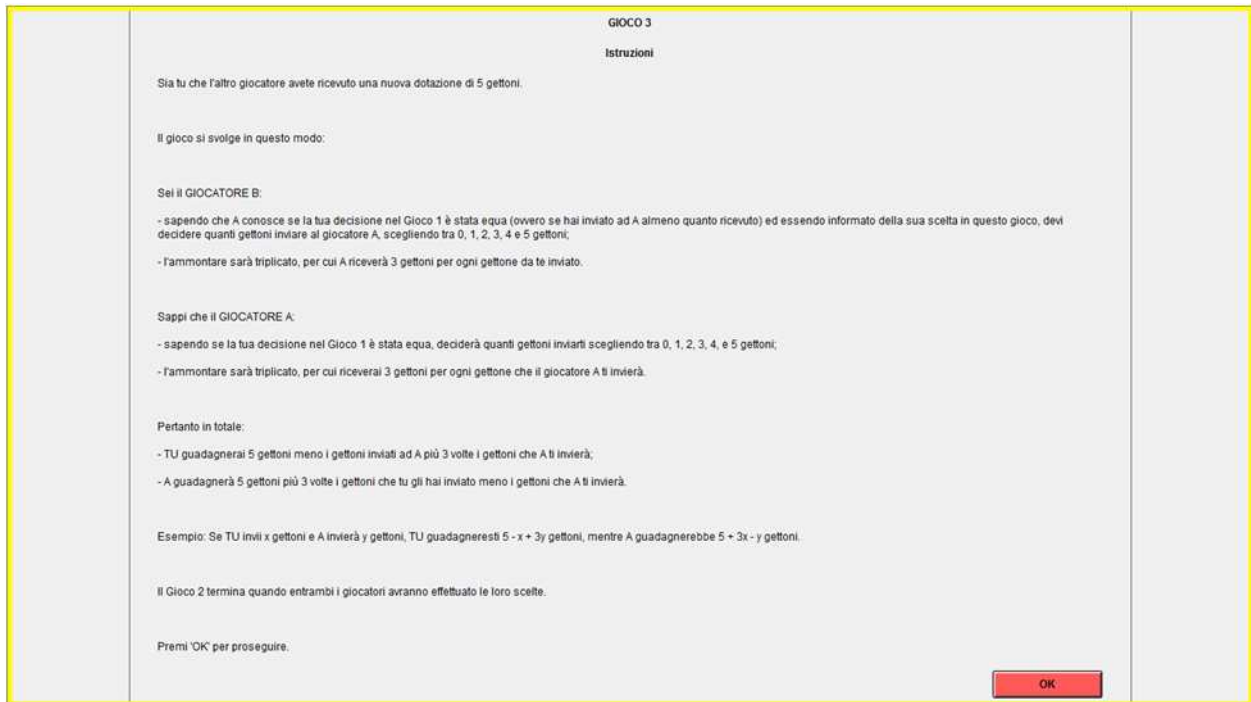
SCREENSHOT 4B



*English translation (top-down):* Game 2 || In this game PLAYER A, knowing that in Game 1 you were unfair, he decide to sent you a number of tokens equal to 5 || How many tokens do you want to send to PLAYER A? || Confirm

*Notes:* If player B in Game 1 sent to A at least as much as they received, the word “unfair” in the above text would have been substituted by “fair.”

## SCREENSHOT 5B



*English translation (top-down):* Game 3 || Instructions || Both you and the other player have received an endowment of 5 tokens || The game goes as follows: || You are PLAYER B || - knowing both that A is aware if your decision in Game 1 was fair (that is, if you send to A at least as much as you received) and their choice in this game, you should decide how many tokens to send to player A by choosing between 0, 1, 2, 3, 4, and 5 tokens; || - the amount will be tripled, thus A will receive 3 tokens per each token you sent. || Notice that player A: || - knowing if your decision in Game 1 was fair, will decide how many tokens to send you by choosing between 0, 1, 2, 3, 4, and 5 tokens; || - the amount will be tripled, thus you will receive 3 tokens per each token A sent you. || Therefore, in total: || YOU will earn 5 tokens minus the tokens sent to A plus 3 times the tokens A will send you || A will earn 5 tokens plus 3 times the tokens you sent them minus the tokens A will send you || For example: if YOU send  $x$  tokens and A will send you  $y$  tokens, YOU would earn  $5 - x + 3y$  tokens, while A would earn  $5 + 3x - y$  tokens. || Game 3 ends when both players would have made their choices. || Press 'OK' to continue. || OK

SCREENSHOT 6B

**GIOCO 3**

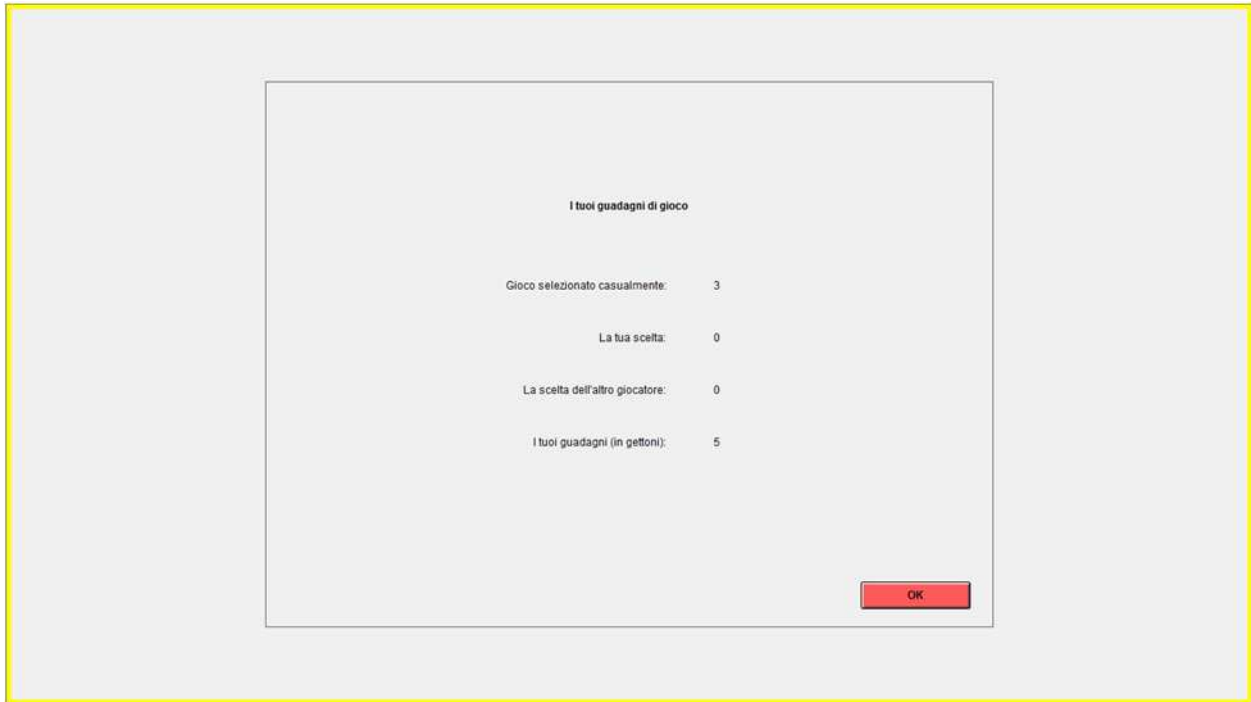
In questo gioco il GIOCATORE A, sapeva che nel Gioco 1 sei stato iniquo.

Rispondi ai seguenti sei quesiti, sapendo che quello corrispondente all'effettiva scelta di A potrebbe concorrere alla determinazione dei tuoi guadagni.

1. Se il giocatore A avesse scelto di inviarti 0 gettoni, TU invieresti ad A un numero di gettoni pari a:	0 1 2 3 4 5 ○ ○ ○ ○ ○ ○
2. Se il giocatore A avesse scelto di inviarti 1 gettoni, TU invieresti ad A un numero di gettoni pari a:	0 1 2 3 4 5 ○ ○ ○ ○ ○ ○
3. Se il giocatore A avesse scelto di inviarti 2 gettoni, TU invieresti ad A un numero di gettoni pari a:	0 1 2 3 4 5 ○ ○ ○ ○ ○ ○
4. Se il giocatore A avesse scelto di inviarti 3 gettoni, TU invieresti ad A un numero di gettoni pari a:	0 1 2 3 4 5 ○ ○ ○ ○ ○ ○
5. Se il giocatore A avesse scelto di inviarti 4 gettoni, TU invieresti ad A un numero di gettoni pari a:	0 1 2 3 4 5 ○ ○ ○ ○ ○ ○
6. Se il giocatore A avesse scelto di inviarti 5 gettoni, TU invieresti ad A un numero di gettoni pari a:	0 1 2 3 4 5 ○ ○ ○ ○ ○ ○

*English translation (top-down):* Game 3 || In this game PLAYER A knew that in Game 1 you were unfair. || Answer the following questions, knowing that the answer corresponding to the actual A's choice could contribute to the determination of your earnings. || 1. If player A would have chosen to send you 0 token, YOU would send to A a number of tokens equal to 0\_ 1\_ 2\_ 3\_ 4\_ 5\_ || 2. If player A would have chosen to send you 1 token, YOU would send to A a number of tokens equal to 0\_ 1\_ 2\_ 3\_ 4\_ 5\_ || 3. If player A would have chosen to send you 2 tokens, YOU would send to A a number of tokens equal to 0\_ 1\_ 2\_ 3\_ 4\_ 5\_ || 4. If player A would have chosen to send you 3 tokens, YOU would send to A a number of tokens equal to 0\_ 1\_ 2\_ 3\_ 4\_ 5\_ || 5. If player A would have chosen to send you 4 tokens, YOU would send to A a number of tokens equal to 0\_ 1\_ 2\_ 3\_ 4\_ 5\_ || 6. If player A would have chosen to send you 5 tokens, YOU would send to A a number of tokens equal to 0\_ 1\_ 2\_ 3\_ 4\_ 5\_ || Confirm

SCREENSHOT 7B



*English translation (top-down):* Your game earnings || Game randomly selected:3 || Your choice: 0 || The choice of the other player: 0 || Your earnings (in tokens): 5 || OK