



DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

TRUST RESPONSIVENESS AND BELIEFS

Gerardo A. Guerra and Daniel John Zizzo

Number 99

March 2002

Manor Road Building, Oxford OX1 3UQ

Trust Responsiveness and Beliefs

Gerardo Guerra

Department of Economics, University of Oxford, Oxford OX1 3UQ, UK.

Daniel John Zizzo*

Department of Economics, University of Oxford, Oxford OX1 3UQ, UK.

(Version April 25, 2002)

Abstract

Trust responsiveness is the tendency to fulfill trust because you believe that it has been placed on you. The experiment presented in this paper uses two simple trust games to measure directly or indirectly the robustness of trust responsiveness in three conditions: when beliefs are elicited and a summary of these beliefs is transmitted; when beliefs are elicited but not transmitted; when beliefs are not elicited. Insofar as we can tell, trust responsiveness is robust to our belief manipulations: this strengthens the case for the real-world significance of trust responsiveness. Shame provides a possible explanation for unexpected trusters' choices.

JEL classification: C72; C91; Z13

Keywords: Trust; Trust responsiveness; Belief elicitation; Shame

* Corresponding author. Email addresses: gerardo.guerra@wolfson.ox.ac.uk (G. Guerra), daniel.zizzo@economics.ox.ac.uk (D.J. Zizzo). We would like to thank Gabriel Goddard, Jahirul Islam, and Angelica Leroux for experimental assistance, Michael Bacharach for guidance, and British Telecom Research Laboratories (under agreement ML826341) and the Senior Tutor's Fund, Wolfson College, Oxford, for financial support. The usual disclaimer applies.

Trust Responsiveness and Beliefs

1. Introduction

This paper presents an experimental test of the empirical robustness of a psychological hypothesis, *trust responsiveness*, that explains why people honor trust. Trust is usually considered of great importance to minimize the costs of writing, policing and enforcing contracts, and thus to enhance efficiency in economic transactions (e.g., Arrow, 1974). It is at the heart of what is usually meant by “social capital” (Paldam, 2000), and, even though this is a controversial notion (Durlauf, 2002), one needs not to subscribe to it to recognize that the lack of trust may be a cause of underdevelopment in realities such as Southern Italy (Putnam, 1993) or transition economies (Miształ, 1996). Further illustrations of where trust, or the lack thereof, plays an essential role include areas as diverse as e-commerce, the financial system, tax payment, and the choice by a customer of a particular supplier when there are switching costs and hold-up problems may occur. The importance of trust is clear, although modifying the degree of trust in a society has proven a difficult task.

In a simple two-player environment, we can call *truster* the agent who places trust and *trustee* the agent who fulfills trust. To say that a truster trusts a trustee means at the very least that the truster believes that the trustee will not exploit a vulnerability that the truster creates by her own action (James, 2002). The paradox of trust is that the trustee does not exploit this vulnerability and so it is on the part of the trustee rather than of the truster: this is because typically, from the viewpoint of the truster, trust is the best response to an expectation of trust fulfilment (Hausman, 1998).

Trust responsiveness is a possible psychological mechanism that explains why trustees fulfill trust, which solves (at least partially) the paradox of trust: it is a tendency to fulfill trust because you believe that it has been placed on you (Bacharach et al., 2001), perhaps out of the feeling of not wanting to “let down” the truster (Dufwenberg and Gneezy, 2000). A trust responsive agent is more likely to fulfill trust the higher her assessment of the probability (the more she believes) that she is being trusted by her coplayer. Consider the truster’s first-order belief f^* that the trustee will fulfill the trust, and the trustee’s second-order belief f^{**} of the truster’s first-order belief: then trust responsiveness predicts a correlation between f^{**} and trust fulfilment.

The importance of proving both the existence and robustness of trust responsiveness arises from the influence it can exert on trust, therefore providing an instrument to improve trust-based efficiencies in society.

The experiment presented in this paper uses two simple trust games to measure (directly or indirectly) the robustness of trust responsiveness under three experimental conditions: when beliefs are elicited and a summary statistics of these beliefs is transmitted; when beliefs are elicited but not transmitted; when beliefs are not elicited. We find that, insofar as inferences can be made, trust responsiveness seems surprisingly robust to different belief manipulations regimes. This strengthens the case for trust responsiveness as an explanation for how and when the trust paradox is successfully solved in the real world.

We also mostly find support for the idea that, since trust is usually the best response to an expectation of trust fulfilment, we should expect a positive correlation between trusting and f^* ; however, we do not find a significant correlation in one of the two games when beliefs are transmitted. The idea of shame may help to explain this finding, though other stories are undoubtedly possible.

The rest of this paper is organized as follows. Section 2 describes the background to the experiment. Sections 3 and 4 present the design and results, respectively. Section 5 contains a discussion on the nature of trust responsiveness and on the results, while section 6 concludes. The appendices contain the experimental instructions and the raw data.

2. Experimental Background

A number of experimental papers on games that present scope for trust have appeared in recent years (for example, Berg et al., 1995; Abbink et al., 2000; Guerra, 2002 McCabe et al., 2002). The Prisoner's Dilemma (PD), and other symmetric social dilemmas, are cases where there is scope for trust on the part of all players, and as such they are often interpreted as games involving trust (e.g., James, 2002). In this context it is well known that cooperation increases in the expectation of cooperation by the other players (e.g., Croson, 2000): this is consistent with trust responsiveness, although it is also consistent with more standard explanations, such as kindness (Rabin, 1993), inequality aversion (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000) or kindness-reciprocity (Falk and Fischbacher, 2001).

Bacharach et al. (2001) focus on a general class of 2×2 asymmetric games that they label Basic Trust Games or BTG: the payoff matrix of the BTG is shown in Figure 1 and is characterized by the following essential inequalities: 1) $y < a$ (Exposure); 2) $a < w$ (Improvement); 3) $x < z$ (Temptation).

(Insert Figure 1 about here).

The Trust-Honor Game of James (2002) satisfies these essential inequalities, although it varies the b payoff obtained by the trustee if the truster withholds trust, according to whether the truster decides to fulfill or violate trust. The Nash equilibrium of the game for rational self-interested agents is clearly *(Withhold, Violate)*, whereas *(Trust, Fulfill)* is the outcome where trust is placed and fulfilled. In the context of the BTG, f^* is the probability assigned by the truster that the trustee will play *Fulfill* (F), and f^{**} is the trustee's estimate of f^* : trust responsiveness then predicts a positive correlation between F and f^{**} . Bacharach et al. (2001, section 1.5) prove that in the BTG inequality aversion and Rabin kindness do not predict this correlation, while Fehr and Fischbacher's model actually predicts a negative correlation. This last result is somewhat counterintuitive and a brief explanation may be due (for the analytical details, see Bacharach et al., 2001). The greater f^{**} is, the more a trustee motivated by kindness-reciprocity believes that the truster is being unkind to him: this is because, from the viewpoint of the trustee, a truster who believes that F is more likely is one that by playing Trust is imposing a lower expected payoff on the trustee relative to the truster (since $w > y$ by the Exposure and Improvement conditions and $x < z$ by the Temptation condition), therefore behaving in an unkindly manner.

Bacharach et al. (2001) consider three types of BTG, and the two of them that are of direct interest for our paper are the Kind Trust Game (KTG) and the Needy Trust Game (NTG), also displayed in Figure 1. The KTG is a 2×2 normal-form variant of the kind of asymmetric trust games that have been common in the literature (e.g., Dufwenberg and Gneezy, 2000), where trusting is a "kind" act to the extent that it gives the possibility of a higher payoff to the trustee than if the trustee did not trust. In addition, the NTG makes the truster particularly "needy" of trusting, since if she does not trust she gets a comparatively bad outcome. Of course other reasons, quite independent of need, may lead the truster to trust more in the NTG than in the KTG, such as an egalitarian or utilitarian motivation. Conversely, McCabe et al. (2002) argue that a reduction in the opportunity cost of trusting will make the trustee less likely to fulfill, as the trustee will perceive the act of trust

as less kind: if this effect were to hold, we should expect a lower fulfilling rate in the NTG than in the KTG, and, as a consequence, less trusting in the NTG than in the KTG.

In Bacharach et al.'s data (2001), 0.52 and 0.61 of the trusters played *Trust* in the KTG and NTG, respectively, and the corresponding F rates were 0.40 and 0.52. Beliefs were elicited in an incentive-compatible way on the part both of trusters and of trustees. In every round each truster was asked to state as a probability her confidence that the trustee would fulfill her trust. Each trustee was told the average value of the probability assessments of the trusters he was not matched with for any given round. Trustees were then asked to state their confidence that they were being trusted by their coplayer. Subjects would then play the trust games in their normal form, choosing actions *simultaneously*.

Significant evidence for trust responsiveness was found. By having belief elicitation and transmission of beliefs from trusters to trustees, though, results may have not been very general. Belief transmission was motivated in Bacharach et al. (2001) by the desire to enable trustees to form definite enough beliefs about the trusters' confidence in them. While this is reasonable, we cannot exclude the possibility that this may have biased trust responsiveness upwards for at least two reasons: first, because it may have drawn the attention of subjects to the truster's confidence in them, and this may have made any psychological factor leading to trust responsiveness more salient than it otherwise would have been; second, because it is not obvious that in the real world people typically receive this kind of information.

Dufwenberg and Gneezy (2000) did not have any transmission of beliefs from trusters to trustees. Trusters chose whether to take a sum of money $k < 20$ guilders or to leave it and have the trustees choose how to reward them, by an amount $j \in [0, 20)$. Beliefs were elicited from trusters and trustees: trusters were asked to guess the average j of all trustees incentive-compatibly – call this guess j^* –, while trustees were asked to guess the j^* of the trusters who placed trust in them – call this second-order guess j^{**} -. A positive correlation between j and j^{**} was found.

Dufwenberg and Gneezy's (2000) results broadly suggest that trust responsiveness should hold even if belief transmission did not take place. However, they are far from decisive: trust responsiveness is a hypothesis relating trust fulfillment to a probability, f^{**} , that trust is placed, and the relationship between j^{**} (the guess of an amount of money) and f^{**} (a probability) is not entirely clear. Since in Dufwenberg and Gneezy fulfillment is a many-valued variable, j^* is not an

unambiguous measure of f^* : j^* is the guess of the average sum of money Ej rather than of a probability, a guess compatible with many subjective probability distributions over the support of j (Bacharach et al., 2001). This ambiguity infects the corresponding j^{**} estimate: it is not clear that the trustee should behave in the same way for a given j^{**} , whatever the underlying subjective distribution of j values that lie behind it.

A more general problem both with Dufwenberg and Gneezy (2000) and Bacharach et al. (2001) lies in the very fact that beliefs are elicited. This is a problem that can be dealt only partially in an experimental setting: belief elicitation is essential in order to measure trust responsiveness in an experimental setting, and so one cannot test directly and conclusively whether the lack of belief elicitation would imply a different value of trust responsiveness. Nevertheless, there is at least some cause of concern: Croson (2000) discusses how belief elicitation distorts behavior in the context of social public goods and PD experiments.

If trusting or fulfilling rates were significantly different with and without belief elicitation, then there would be also a concrete worry that belief elicitation distorts behavior in trust games. This would cast a shadow on the actual degree of relevance of trust responsiveness in real-world contexts without belief elicitation. Hence, an indirect and admittedly partial test of the robustness of trust responsiveness when beliefs are not elicited is to check that observable measures of trusting and fulfilling do not change in comparable conditions with and without belief elicitation.

This paper presents the results of an experiment that addresses these problems. We compare the outcomes of three experimental treatments employing the KTG and NTG in counterbalanced order: condition 1 with no belief elicitation, condition 2 with belief elicitation but no belief transmission, and condition 3 (from Bacharach et al., 2001) with both. We cannot obviously have a condition with belief transmission but no belief elicitation, as beliefs need to be elicited first in order to be transmitted. We can think of the experimental design as one that allows an increasing institutional role for beliefs as one moves from condition 1 to condition 2 and from condition 2 to condition 3. In employing two BTG variants, we ensure that f^* and f^{**} are measurable entities when beliefs are elicited, in conditions 2 and 3. In moving from condition 3 to condition 2, we can then check whether trust responsiveness remains unaltered or disappears. In moving from condition 2 to condition 1, we can check whether the measures that are observable in both conditions, i.e. trusting and fulfilling rates, change or otherwise. Thus, we can have in the first case a clean and

direct test, and in the second case a partial and indirect test, of the robustness of trust responsiveness.

3. Design

The experiment was run in the Department of Economics in the University of Oxford in February, November and December 2001. Recruitment was done by printed and email advertisements, and most subjects were undergraduate or postgraduate students. Eight subjects participated to each session, playing four trust games each either as trusters or as trustees, and keeping the same role throughout the experiment. There were ten experimental sessions, two in condition 1 and four each in conditions 2 and 3. The design thus was meant to produce 32 observations for each of the two roles in condition 1, and twice as many in conditions 2 and 3 (we had extra power in conditions 2 and 3 because of the need of measuring the trust responsiveness coefficient). The BTG used were the KTG or NTG games illustrated in Figure 1: subjects played first two rounds of one game and then two rounds of the other game. The order of the presentation of the games was counterbalanced within each condition (although already Bacharach et al., 2001, found no order effects). The experiment was fully computerized, and the instructions can be found in Appendix A.

A session consisted of three stages. At the start of it, subjects were assigned randomly to terminals separated by screens. In the Introduction Stage, the nature of the tasks and the payment procedure were explained, with examples and practice; subjects were also assigned a “starting credit” of 4 U.K. pounds. Next, four subjects were assigned randomly to the role of either trusters or trustees (here and elsewhere, of course, the terminology was neutral in the instructions).

The Play Stage contained the core of the experiment. In each of the four rounds the order of events was as follows:

1. *In all conditions*, each subject was shown the payoff matrix of the game to be played (either a KTG or a NTG), in the form of a “points table”.
2. *In conditions 2 and 3 only*, each truster made a *statement* of the probability she attached to the event that her coplayer would choose strategy *F*.
3. *In condition 3 only*, each trustee received a *report* consisting of the mean value of the statements of his non-coplayers.

4. *In conditions 2 and 3 only*, each trustee made a *guess* at the statement of his coplayer.

5. *In all conditions*, each player made her BTG strategy choice simultaneously.

In relation to conditions 2 and 3, the statement measured the truster's confidence f^* , and the guess measured the trustee's belief about his coplayer's statement and so measured his confidence-perception f^{**} . Statements and guesses were made by using the mouse to manipulate a pointer on a semicircular dial calibrated in integers from 0 to 100. Reports were rounded to the nearest integer.

Subjects were told nothing at the end of rounds 1, 2 or 3 about the strategy choices of others in either the current or earlier rounds. In conditions 1 and 2 they received no feedback at all; in condition 3, the only information any subject received about other subjects' behavior was the report about trusters' statements given to trustees.

In the Payment Stage, subjects were paid: a) the starting credit of £ 4; b) the strategy payment; c) in conditions 2 and 3 only, the guess payment.

A round was randomly chosen for the strategy payment, and subjects were paid for the strategy choices they had made then, in accordance with the game payoff matrix that they had played in that round. Strategy payments could be negative, leading to the loss of part of the starting credit. They ranged from £ -1.5 to £ 4.5.

In conditions 2 and 3 a different round was also randomly chosen to determine the guess payment: trusters were paid for their statements and trustees for their guesses in this round, according to the same procedure as Bacharach et al. (2001). Statements were paid according to the well-known quadratic scoring rule (Davis and Holt 1993). A truster stating s received £ $3[1-(1-0.01s)^2]$ if her coplayer chose F and £ $3(1-0.01s)^2$ if she did not. Guesses were paid according to a triangular scheme (Croson 2000): a trustee guessing exactly correctly received £ 3 if the guess was exactly correct, and 30 pence less for each unit of error, subject to non-negativity of the payment (so she received nothing if his guess was 10 or more percentage points out).

Average payments were £5.84 for about 50-55 minutes of work in conditions 2 and 3; condition 1 subjects earned an average £5.41 for about 35-40 minutes of work.

4. Results

4.1 Trusting and fulfilling rates

The raw experimental data can be found in Appendix B. Table 1 shows observed trusting and fulfilling rates by condition and game. Trusting rates were 0.688, 0.500 and 0.633 in conditions 1, 2 and 3, respectively; the corresponding fulfilling rates were 0.375, 0.328 and 0.484.

(Insert Table 1 about here).

F tests are insignificant in relation to condition (for trusting: $F = 1.946$, $d.f. = 2$, $P > 0.1$; for fulfilling: $F = 1.675$, $d.f. = 2$, $P > 0.1$), to game (for trusting: $F = 2.353$, $d.f. = 1$, $P > 0.1$; for fulfilling, $F = 0$, $d.f. = 1$, $P > 0.1$), and to the interaction effect (for trusting: $F = 0.235$, $d.f. = 2$, $P > 0.1$; for fulfilling: $F = 1.052$, $d.f. = 2$, $P > 0.1$). Two t tests are only marginally insignificant: a t test comparing the trusting rates in condition 2 against the other two gives $t = 1.911$, $d.f. = 154$, $P = 0.058$ (all t tests here and below are two-tailed); a t test comparing the fulfilling rates in conditions 2 and 3 gives $t = 1.809$, $d.f. = 126$, $P = 0.073$. However, one needs to be cautious in considering significance because these tests implicitly assume that the decisions made by each subjects are independent observations, which might not be a reasonable assumption: if not, the tests would tend to overstate significance.

(Insert Figures 2 and 3 about here).

Figures 2 and 3 depict the average trusting and fulfilling rates by experimental condition and game. It can be observed that, while there is no trend across conditions in the KTG, the NTG seems to result in a fall in the fulfilling rate in moving from condition 3 to condition 2 ($t = 2.269$, $d.f. = 78$, $P < 0.05$), but there is no difference between conditions 1 and 2.

4.2 Statements and guesses

This data was collected in conditions 2 and 3. The average statement by trusters was 33.906 in condition 2 and 39.283 in condition 3: the difference is not significant ($t = 0.997$, $d.f. = 122$, $P > 0.1$). In the KTG, the average statement was 30.656 and 43.433 in conditions 2 and 3 respectively ($t = 1.692$, $d.f. = 60$, $P = 0.096$); in the NTG, the average statement was 37.156 and 35.133 in conditions 2 and 3, respectively ($t = 0.262$, $d.f. = 60$, $P > 0.1$).

The average guess by trustees was 33.500 and 48.500 in conditions 2 and 3, respectively: the difference is significant even at a conservative level ($t = 3.403$, $d.f. = 126$, $P = 0.001$). The KTG average guess was 31.094 and 53.063 in conditions 2 and 3, respectively ($t = 1.235$, $d.f. = 62$, $P >$

0.1); the NTG average guesses were 35.906 and 43.938 in conditions 2 and 3, respectively ($t = 3.716$, $d.f. = 62$, $P < 0.001$). The difference so appears to be larger in the KTG than in the NTG.

4.3 Trust responsiveness

We can measure the role of trust responsiveness in conditions 2 and 3. Define the variable $F = 1$ when a trustee fulfills trust, and equal to 0 otherwise. A plot of the mean F rate against the guess can be used to inspect whether the propensity to fulfill increased with the expectation of being trusted.

(Insert Figure 3 about here).

An inspection of Figure 3 shows that the relative frequency of F rises with the guess, as trust responsiveness would predict. Trust responsiveness requires that trustees who have higher guesses will be more likely to fulfill trust: thus, we would expect the mean guess of trustees who fulfill trust to be higher than that of trustees who withdraw it. In Condition 2 the mean guess was 24.674 when $F = 0$, and 51.571 when $F = 1$, more than double in magnitude ($t = 5.520$, $d.f.=62$, $P < 0.001$); in Condition 3, the means were 38.091 when $F = 0$ and 59.581 when $F = 1$, an increase in magnitude in excess of 60% ($t = 3.481$, $d.f.=62$, $P = 0.001$). The Pearson correlation between F and guess was 0.404 ($P = 0.001$) in condition 3, and increases to 0.555 in condition 2 ($P < 0.001$). A similar pattern holds if one uses a non-parametric measure: the Spearman correlations between F and guess were 0.521 ($P < 0.001$) and 0.394 ($P = 0.001$) in conditions 2 and 3, respectively. It is not the case that there was less trust responsiveness in condition 2 than in condition 3.

Regression analysis can be performed to estimate the coefficients of trust responsiveness in the two conditions. Let f be the propensity to fulfill, i.e the probability that the trustee plays F . Table 2 contains the results of the estimation of probit models of the form:

$$f = \Phi(\beta \cdot x) \tag{1}$$

where $\Phi(y)$ denotes the probability that a standard normal variate is less than y .

(Insert Table 2 about here).

The explanatory variables are of two types: they either refer to level effects that work independently of trust responsiveness (Ntg , equal to 1 when the NTG is played and 0 otherwise, and

Communication, equal to 1 in condition 3 and equal to 0 in condition 2 or they refer to guess interaction effects that measure trust responsiveness coefficients and allow them to vary according to the condition and game employed. The standard errors are corrected to control for individual-specific effects. As in Bacharach et al. (2001), the level variables are insignificant. The only interaction variable involving *Communication* and which is only marginally significant is *Communication x Guess*, suggesting *less* trust responsiveness with than without communication. However, the joint elimination of all variables except *Ktg x Guess* (where *Ktg* is equal to 1 when the KTG is played and 0 otherwise) and *Ntg x Guess* is accepted using a Wald test [$\chi^2(4) = 3.61, P > 0.1$]. The same outcome is achieved if an iterative procedure is used, for example if one first eliminates the two least significant variables *Ktg* and *Communication* [$\chi^2(2) = 1.61, P > 0.1$], and only afterwards one removes the remaining two [$\chi^2(2) = 1.52, P > 0.1$].

4.4 Other results

We checked for the role of a few demographical variables. Economics students tended to trust less [$r(\text{Economics dummy, trusting choice}) = -0.347, P < 0.001$], while older people tended to trust more [$r(\text{age, trusting choice}) = 0.282, P < 0.001$]. We found no correlation between demographical variables and fulfilling trust.

An intriguing and slightly puzzling finding emerges by looking at the correlation between statements and choices by trusters (where *Trust* = 1 if a truster trusted and 0 otherwise). In relation to the KTG, the Pearson correlation is virtually identical between condition 2 and 3 [$r(\text{statement, Trust}) = 0.497$ and 0.490 , respectively, with $P < 0.001$]. However, in relation to the NTG, while in condition 2 we have $r(\text{statement, Trust}) = 0.533$ ($P < 0.001$), which is comparable to the KTG values, in condition 3 we have $r(\text{statement, Trust}) = 0.208$, which is not significant.

5. Discussion

Trust responsiveness is a psychological mechanism the individuation of which probably dates back to at least David Hume's *Treatise*: in his view, "sympathy" carries so far as to make us internalise other people's moral judgements and act in order "not only to be loved and praised, but also to be worthy of love and praise" (Brown, 1994, p. 30). If a trustee is expected (trusted) by the truster to follow a social norm of fulfilling trust, then "being worthy of love and praise" implies an

obligation to meet that expectation (fulfill that trust). If he does not, the truster will be disappointed with the trustee, and, if the trustee's utility depends on the truster's opinion of him because of Humean "sympathy", he will be trust responsive. This is a theory of trust responsiveness based on "person disappointment", and has appeared in various guises in the work of economists (Sugden, 1998), philosophers (Pettit, 1995), psychologists (Jussim, 1986) and sociologists (Granovetter, 1985). A different theory could be hypothesized based on "outcome disappointment": the trustee may be concerned of letting down the truster because by doing so the truster may be disappointed of the outcome she gets. A combination of regret or disappointment aversion on the part of the truster (Jia et al., 2001) and altruism on the part of the trustee will then be sufficient to produce trust responsiveness, as the trustee will be less keen to withhold trust the more he believes the truster is trusting him to get him to the "good" outcome.

The main objective of this paper was to check the robustness of trust responsiveness to alternative schemes of belief manipulation. A direct test of this can be achieved in moving from condition 3, with belief transmission, to condition 2, without belief transmission. The evidence strongly supports the claim that trust responsiveness is no less when trustees do *not* receive any reports from trusters about their f^* than when they do.

A more indirect test has to be used in moving from condition 2 to condition 1, that is from belief elicitation to no belief elicitation. The reason is that trust responsiveness is a relationship between beliefs and behavior, and so cannot be measured if beliefs are not elicited. Nevertheless, we can do a minimal test consisting in checking whether trusting or fulfilling behavior changes in moving from condition 2 to condition 1: if it did, we would have a *prima facie* argument that such behavior in condition 1 is qualitatively different, at least in part, from when beliefs are elicited. While there is some evidence that in the NTG the fulfilling rate decreases in moving from condition 3 to condition 2, there is no such difference, or any other difference, in moving from condition 2 to condition 1, the comparison that matters in relation to our indirect test. Of course, one needs to recognize that the power of our indirect test is limited, but it is still the best that can be made in the light of our observability problem.

We conclude that, to the extent to which inferences can be made, trust responsiveness is robust to alternative schemes of belief elicitation: whether no beliefs are elicited at all, or beliefs are elicited but not transmitted from trusters to trustees, or beliefs are elicited and transmitted. Since

belief elicitation, let alone transmission, may often not occur outside the experimental laboratory, this finding of robustness is a potentially important one: it strengthens the case for the relevance of trust responsiveness in explaining why and when people fulfill trust. The key to getting to the cooperative outcome in trust games is usually considered trust fulfilment, because typically, and certainly in the KTG and NTG, trust is the best response to an expectation of trust fulfilment. Thus, an explanation for trust fulfilment is bound to be a crucial step in predicting when the trust paradox can be solved. While we make no claim that trust responsiveness is the *only* factor that matters, we suggest that it has a role to play.

It will be recalled that the statement is the elicited belief by the truster that trust will be fulfilled. The positive correlation between statement and trusting choice in condition 2 with the KTG and in condition 3 with both KTG and NTG, endorses the view that, the more trusters expect trustees to fulfill trust, the more we should expect trusters to trust. The insignificance of the correlation for NTG play in condition 2 is slightly puzzling, but the correlation is still correctly signed. If, however, we were to take the insignificance of the correlation at its face value, a possible interpretation may be that, in the NTG, subjects are ashamed to appear to want to rely on other players, and this produces a wedge between the true f^* and the statement to be transmitted to the trustees as part of a summary statistic. Since behavior is driven by the true f^* rather than by the distorted f^* , we would then expect a lower correlation between statement and trusty choice in the NTG, as we find. This “shame” conjecture clearly requires proper experimental testing, but has a real-world parallel in the social stigma associated to the receipt of welfare benefits by subjects who are legally entitled because of need (e.g., Besley and Coate, 1992). For example, according to Riphahn (2001) more than half of German households eligible for social assistance do not claim their benefits. According to this conjecture, this is largely because of the stigma associated to reporting their need.

Trust responsiveness and shame are motivations intimately related to the beliefs of others and, as such, they should best be formalized within the framework of psychological games (Geanakoplos et al., 1989).

6. Conclusions

Trust responsiveness is a possible psychological mechanism for why trustees fulfill trust, and hence a solution to the paradox of trust: it is a tendency to fulfil trust because you believe that it has been placed on you. The experiment presented in this paper uses two simple trust games to measure the robustness of trust responsiveness under three experimental conditions: when beliefs are elicited and a summary statistic of trusters' first-order beliefs is transmitted to trustees; when beliefs are elicited but not transmitted; when beliefs are not elicited. In moving from belief transmission to no belief transmission, we found no decrease in trust responsiveness. In further moving to no belief elicitation, given that we cannot measure trust responsiveness without the measurement of beliefs, an indirect and partial test is to check whether there is any change in observable variables such as trusting and fulfilling choices, and we find none. We conclude that, insofar as we can tell (a qualification required to take into account of the tentativeness of our indirect test), trust responsiveness is robust to whether beliefs are elicited and transmitted or otherwise. This strengthens the case for the real-world importance of trust responsiveness.

Trust responsiveness is an explanation of why trustees fulfill trust, and this is the key to the paradox of trust whenever, as in our games, trust is the best response to an expectation of trust fulfillment. We find that the correlation between trusting choices and expectation of trust fulfillment is indeed positive, and mostly significantly so; we advance the conjecture that shame on the part of trusters who do not want to appear needy may explain the one anomalous case we find on this, and may have a real-world parallel in the social stigma of claiming welfare benefits.

References

- Abbink, K., Irlenbusch, B., Renner, E., 2000. The moonlighting game: an experimental study on reciprocity and retribution, *Journal of Economic Behavior and Organization* 42, 265-277.
- Arrow, K., 1974. *The Limits of Organization* (Norton, New York).
- Bacharach, M., Guerra, G., Zizzo, D.J., 2001. Is trust self-fulfilling? An experimental study. Discussion Paper n. 76 (October), Department of Economics, University of Oxford.
- Berg, J., Dickaut, J., McCabe, K., 1995. Trust, reciprocity and social history. *Games and Economic Behavior* 10, 122-142.
- Besley, T., Coate, S., 1992. Understanding welfare stigma: taxpayer resentment and statistical discrimination. *Journal of Public Economics* 48, 165-183.

- Bolton, G.E., Ockenfels, A., 2000. ERC: a theory of equity, reciprocity, and competition. *American Economic Review* 90, 166-193.
- Brown, C., 1994. From spectator to agent: Hume's theory of obligation. *Hume Studies* 20, 19-31.
- Croson, R.T.A., 2000. Thinking like a game theorist: factors affecting the frequency of equilibrium play. *Journal of Economic Behavior and Organization* 41, 299-314.
- Davis, D., Holt, C., 1993. *Experimental Economics*. Princeton University Press, Princeton.
- Dufwenberg, M., Gneezy, U., 2000. Measuring beliefs in an experimental lost wallet game. *Games and Economic Behavior* 30, 163-182.
- Durlauf, S.N., 2002. Bowling Alone: a review essay, *Journal of Economic Behavior and Organization* 47, 259-273.
- Falk, A., Fischbacher, U., 2001. Distributional consequences and intentions in a model of reciprocity. *Annales d'Economie et de Statistique* 63-64, 111-129.
- Fehr, E., and Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114, 817-868.
- Geanakoplos, J., Pearce, D., Stacchetti, E., 1989. Psychological games and sequential rationality, *Games and Economic Behavior* 1, 60-79.
- Granovetter, M., 1985. Economic action and social structure: the problem of embeddedness. *American Journal of Sociology* 91, 481-510.
- Guerra, G., 2002. Crowding out trust: the adverse effects of verification. An experiment. Discussion Paper n. 98 (April), Department of Economics, University of Oxford.
- Hausman, D., 1998. Fairness and trust in game theory. Unpublished paper, London School of Economics and University of Wisconsin.
- James Jr., H.S., 2002. The trust paradox: a survey of economic inquiries into the nature of trust and trustworthiness. *Journal of Economic Behavior and Organization* 47, 291-307.
- Jia, J., Dyer, D.S., Butler, J.C., 2001. Generalized disappointment models. *Journal of Risk and Uncertainty* 22, 59-78.
- Jussim, L., 1986. Self-fulfilling prophecies: a theoretical and integrative review. *Psychological Review* 93, 429-445.
- McCabe, K.A., Rigdon, M., Smith, V.L., 2002. Positive reciprocity and intentions in trust games. *Journal of Economic Behavior and Organization*, in press.

- Misztal, B., 1996. *Trust in Modern Societies: The Search for the Basis of Social Order* (Polity Press, Cambridge).
- Paldam, M., 2000. Social capital: one or many? Definition and measurement. *Journal of Economic Surveys* 14, 629-653.
- Pettit, P., 1995. The cunning of trust. *Philosophy and Public Affairs* 29, 202-225.
- Putnam, R., 1993. *Making Democracy Work: Civic Traditions in Modern Italy* (Princeton University Press, Princeton).
- Riphahn, R.T., 2001. Rational poverty or poor rationality? The take-up study of social assistance benefits. *Review of Income and Wealth* 47, 379-398.
- Sugden, R., 1998. Normative expectations: the simultaneous evolution of institutions and norms. In: A. Ben-Ner and L. Putternam (Eds.), *Economics, Values, and Organization*, Cambridge University Press, Cambridge, pp. 73-100.

Appendix A. Experimental instructions

The instructions are divided according to the stage (Introduction Stage, Play Stage and Payment Stage) and the role (trustees, or R players, and trustees, or E players). Sentences specific to a condition are placed in bold square brackets, preceded by the number of the condition, also in bold.

INTRODUCTION STAGE: R AND E PLAYERS

Welcome to the Department of Economics. You are about to take part in an experimental study of decision making.

You are not allowed to speak to other participants or communicate in any other way. If you want to ask a question, please put up your hand.

The experiment is in three stages, called the INTRODUCTION, STAGE 1 and STAGE 2. In Stages 1 and 2 you will be asked to make some decisions.

In the Introduction stage, we will explain the nature of the decisions and give you practice.

As of this moment, you have been credited with the sum of £4, your STARTING CREDIT. During the course of the experiment you may, through the decisions you make, either add to this, leave it unchanged, or lose some of it. The net amount to your credit at the end of the experiment will be paid to you before you leave the building.

In the Introduction stage, we will explain the nature of the decisions and give you practice.

Please wait for other participants.

INTRODUCTION

There [1 is one] [2 & 3 are three] kinds of decisions in this experiment: INTERACTIVE DECISIONS [2 & 3 , LIKELIHOOD DECISIONS and GUESSING DECISIONS]. We will call them I DECISIONS [2 & 3 , L DECISIONS and G DECISIONS] for short.

Everyone will make some I DECISIONS [2 & 3 , and each of you will make either some L DECISIONS or some G DECISIONS].

None of your decisions will be revealed by us to other participants either during the experiment or after it.

In your I Decisions (I for Interactive) you will be paired with another participant, your OPPOSITE NUMBER in that decision. In each I Decision you will have a different Opposite Number. You will not be told who your Opposite Number is, either during the I Decision or later. Nor will your Opposite Number know that it is you that she/he is interacting with.

We'll now explain the decisions you will make in more detail, give you some practice on them, and ask some questions to make sure everything is clear to you. We begin with I Decisions.

At any stage of the explanation, and throughout the experiment, you will be able to return to a HELP page which summarizes what you have been told so far, by clicking the button marked HELP.

I DECISIONS

In an I Decision, you choose between two OPTIONS, and your Opposite Number also chooses between two Options. Before making your choice you see a POINT TABLE like the one above. In it, we have labelled your Options TOP and BOTTOM, and your Opposite Number's LEFT and RIGHT.

The numbers in the table show the number of POINTS you and your Opposite Number would get, for various combinations of choices by you and her/him; the one on the left of the comma, in blue, shows YOUR payoff, and the one on the right of the comma, in grey, shows your Opposite Number's. In this example, if the Option you chose was BOTTOM and the one your Opposite Number chose was RIGHT, you would get 1.5 Points and she/he would get 0.5 Points. Notice that

what you get depends on what you both choose. BOTTOM gets you more points if your Opposite Number chooses RIGHT, but fewer if your Opposite Number chooses LEFT.

At the end of the experiment participants will be paid in money for one of the I Decisions they make in Stages 1 and 2. The more points a participant scores the more she/he will be paid. We will explain how the payment is determined in more detail a little later.

We'll now ask you a few questions to make sure you are clear about what I Decisions are and what Point Tables show. After you have entered your answer, you will be given the correct answer and an explanation. If you need further help, please raise your hand.

(Subjects were presented with the following game:

	LEFT	RIGHT
TOP	1.5, 0.5	1.0, 2.0
BOTTOM	0, 2.5	1.5, 0.5

and asked the following questions: if their answer was wrong, they were told: "FALSE! The correct answer is " and the corresponding answer. If they were right they were told: "TRUE! " and the corresponding answer).

(Question 1) What would you get if you chose TOP and your Opposite Number chose LEFT? Enter a number in the box.

(Answer 1) 1.5, because this is the BLUE number in the TOP row and the LEFT column.

(Question 2) What is the worst possible outcome for you in terms of Points if you chose BOTTOM?

(Answer 2) 0. If you chose BOTTOM you would get 0 if your Opposite Number chose LEFT and 1.5 if she/he chose RIGHT, and 0 is less than 1.5.

(Question 3) What is the worst possible outcome for you in terms of Points if you chose TOP?

(Answer 3) 1. If you chose TOP you would get 1.5 if your Opposite Number chose LEFT and 1 if she/he chose RIGHT, and 1 is less than 1.5.

(Question 4) What is the best possible outcome for your Opposite Number in terms of Points if she/he chose RIGHT?

(Answer 4) 2. If he/she chooses RIGHT he would get 2 if you chose TOP and 0.5 if you chose BOTTOM, and 2 is more than 0.5.

(Question 5) Suppose you think your Opposite Number is equally likely to choose LEFT or RIGHT. Do you expect to get a higher payoff from choosing BOTTOM or TOP?

(Answer 5) TOP. If you choose TOP you think you will get 1.5 and 1 with equal likelihood, so your expectation is 1.25; if you choose BOTTOM it is only 0.75.

Ask for help by putting up your hand if you are puzzled. If you are happy, click on OK. *(Subject moved to next screen).*

[Conditions 2 & 3 only from here to the next mark "□":]

[2 & 3 Next, we explain what LIKELIHOOD DECISIONS (L DECISIONS) are, and give some practice on them. In Stages 1 and 2 you will either make L Decisions yourself, or you will interact with participants who do. In the latter case, your ability to make good I Decisions will depend on understanding how L Decisions are made.

L DECISIONS

In an L Decision (L for Likelihood), you are asked to report your opinion on how likely a certain UNKNOWN FACT is to be true. You do this as a percentage figure, called your REPORT. For example, if your Report is 50 percent, this means you think the chance that the Unknown Fact is true is 1 in 2. Similarly, a Report of 100 means you are certain it is true, and a Report of 0 means you are certain it is not true.

We will now give you a practice example in which the Unknown Fact is a simple fact about the world. (In the main part of the experiment, the Unknown Fact will be which Option your Opposite Number chooses.) Here is the example.

PRACTICE L DECISION 1: How likely do you think it is, on a scale of 0 to 100, that A CARD DRAWN FROM A COMPLETE WELL-SHUFFLED PACK WILL BE A SPADE? Use the mouse to move the pointer on the dial to show your answer. (*Subject gives answer*).

Please Press [CONTINUE].

In Stages 1 and 2 each participant who makes L Decisions will be paid for one of them. This payment is called her/his L-PAYMENT. It may be anything up to £3. The L-Payment is determined as follows. You begin with £3. After you have made your Report, an amount gets deducted from your £3 (your L-DEDUCTION) and you get paid what is left. The L-Deduction is calculated by a formula. If the Unknown Fact turns out to be TRUE, then the higher your Report was, the smaller is the Deduction, and the more you are paid. If it turns out to be FALSE, then the lower your Report was, the smaller is the Deduction and the more you get paid. It is important to realize that YOUR EXPECTED L-PAYMENT IS MAXIMIZED IF YOU REPORT YOUR LIKELIHOODS CAREFULLY AND TRUTHFULLY.

This payment scheme is known as the Quadratic Scoring Scheme. The formula, and a full explanation of why it benefits you to report your likelihood carefully and truthfully, are available on request after the experiment (ask for the Handout).

Here's one more example. This time, we'll ask you to make a decision for 'fictitious money'. Imagine you have been given £3. We'll ask you to make an L Decision, then tell you whether the Unknown Fact in this L Decision is true or not, and the payment you would finish up with on this decision.

PRACTICE L DECISION 2: On a scale of 0 to 100, how likely do you think it is that THE RIVER AMAZON IS MORE THAN 3000 MILES LONG?

Remember, if you think that it's a toss up whether the Amazon is longer than 3000 miles, then you maximize what you expect to receive by a Report of 50%, if you think the likelihood is 75%, you maximize it by a Report of 75%, and so on. (*Subject gives answer*).

The length of the Amazon is 3900 miles. Your Report was (*report amount*). Your deduction is therefore £ (*deduction amount*) and your L-Payment is £ (*payment amount*).

G DECISIONS

In a G Decision (G for Guess), you are asked to make a Guess about what your Opposite Number's Report in an L Decision was.

Let us try out a G Decision. In this one, which is purely for practice, we ask you to imagine that you have an Opposite Number.

Imagine that your Opposite Number has just made the following L Decision: 'On a scale of 0 to 100, how likely do you think it is that THE RIVER AMAZON IS MORE THAN 3000 MILES LONG?'. Now here is your G Decision:

PRACTICE G DECISION: Please make your GUESS what your Opposite Number's Report was. (*Subject gives answer*).

In Stages 1 and 2 each participant who makes G Decisions will be paid for one of them. This payment is called her/his G-PAYMENT. The G-Payment rewards participants for the accuracy of their Guesses.

This is the scheme on which G-Payments are determined. You begin with an initial sum of money, and for every percentage point your guess is 'out' a deduction is made. But you cannot lose more than your initial sum of money.]

[□ *end of section for conditions 2 & 3 only*]

THE I DECISIONS OF STAGES 1 AND 2

We come now to the Interactive or I Decisions that you will make in Stages 1 and 2. They represent a type of I Decision that is very common in real life. The BREB group is active in research into discovering how people make I Decisions of this kind.

In these I Decisions there are two people, a MOVER and a RESPONDER. The Mover can choose either MOVE or PASS. If she/he chooses Move, the Responder can choose between RESPONSE A and RESPONSE B, and which one he/she chooses affects how well or badly off both finish up. If the Mover chooses PASS, the Responder cannot affect the position of either.

Here's an example. Theo has a promising research idea, but lacks the resources to develop it alone. Theo has to decide whether to tell a potential collaborator, Alex, about the idea. If Theo does so, then if Alex decides to collaborate, Theo will benefit. However, once Alex is told the idea, there are ways in which Alex can do better for him/herself by not collaborating. Moreover, in this case, Theo will end up worse off than she/he was originally. In this example, Move is sending the research idea, Pass is not sending it, Response A is collaborating if you are told the idea, and Response B is not collaborating if you are told it.

In the real world we find a great variety of Interactive Decisions which have the form of Mover-Responder problems. In Stages 1 and 2 you will be asked to make choices in two different Mover-Responder problems, one in Stage 1 and one in Stage 2.

You will take your I Decisions for POINTS. At the end of the experiment you will be paid for an I Decision at the rate of £1 per Point. We will explain the payment procedure fully in a minute.

Some of you will be Movers and some Responders. Both Movers and Responders may win Points in their I Decisions, but Movers may also in some circumstances lose Points.

Which role you play will be determined by your own choices in a lottery. We will now proceed to this lottery.

THE LOTTERY. You will see a display of 10 'nonsense syllables'. The computer has assigned each of them, randomly, a code number between 1 and 80. No two code numbers are the same. You will be asked to choose one syllable. When everyone has made a selection, the participants with the four lowest code numbers will be assigned the Mover's role and the others the Responder's role. You will be shown immediately to which role you have been assigned.

Please keep your role to yourself both during and after the experiment.

In each of your four I Decisions your Opposite Number will be a different person. You will not learn at any stage who your four Opposite Numbers were.

Please choose one 'nonsense syllable'. (*Subject makes choice*).

Please wait for your Role Assignment. (*Role is assigned*).

PLAY STAGE: R PLAYERS

You have been assigned the Mover role.

STAGE 1: GENERAL DESCRIPTION

This Stage has two ROUNDS. In each Round you face the same Mover-Responder problem, Problem 1. In each Round you will make an I Decision with a different Opposite Number. Your Point Table for Problem 1 is shown above.

Your I Decision is to choose between Move and Pass, and your Opposite Number's is to choose between Response A and Response B. [2 & 3 But before you make your I Decision, you will be asked to make an L Decision. The L Decision is to decide on a Report, on the usual dial, of how likely you think it is that your Opposite Number will choose Response A.]

[3 In thinking about this, you may like to consider what the procedure will be for people in the Responder role. Each Responder, before making his/her I Decision between Response A and Response B, will be given some information about the L Decisions of participants in the Mover role. However, this information will not include any about YOUR L Decision. Instead, your Opposite Number will be told the average Report of the three Movers other than you. For example, if you report the likelihood R1 and the other Movers report the numbers R2, R3 and R4, your Opposite Number will be told the average of R2, R3 and R4, that is, $(R2 + R3 + R4)/3$. (Your Report, R1, will be an ingredient in the average figures given to the Responders who are NOT your Opposite Number.)]

To summarize, the procedure for you in each Round is as follows.

- [2 & 3 • You make your L Decision, a Report about the how likely you think it is that your Opposite Number will choose Response A.]
- [3 • Your Opposite Number is told the average of the L Decisions of all Movers except you.]
- You make your I Decision between Move and Pass. At the same time as this your Opposite Number is asked to make his/her I Decision between Response A and Response B.

Nobody will be told anything about anyone else's I Decisions until the end of the experiment. So you will learn nothing about outcomes of your I Decisions [2 & 3 or of your L Decisions] until the end.

Like Stage 1, Stage 2 consists of two Rounds, in each of which you make [2 & 3 an L Decision and] an I Decision. In each Round you will have an Opposite Number with whom you have not previously interacted in the experiment. The only difference from Stage 1 is in the Mover-Responder problem, which is a different variant, Problem 2, with a different Point Table.

At the end of the experiment your total payment will be determined as follows. The computer programme will randomly choose one of the four rounds of Stages 1 and 2 as your [2 & 3 L-PAYMENT ROUND, and a different round as your] I-PAYMENT ROUND. [2 & 3 You will be reminded of your L Decision in your L-Payment Round and shown the actual choice of your Opposite Number in that round and your resulting L-Payment.] You will be reminded of your I Decision in your I-Payment Round and shown the choice of your Opposite Number in that round and the number of Points you scored, which may be positive or negative. Your final payment will be the sum of

- your Starting Credit of £4,
- [2 & 3 • your L-Payment in your L-Payment Round,]
- the Points you scored in your I Decision in your I-Payment Round, which may be positive or negative, converted into money at £1 per Point.

When everyone is ready, we will begin Round 1 of Stage 1. You may have to wait one or two minutes for others to be ready. We ask you to be patient. Be sure you have understood the whole procedure, referring to Help if you wish to, before clicking on Continue. Put up your hand if you need any further help.

(Loop here after Round 1:)

[2 & 3 (Screen after waiting period for Round 1:)

Please look at the Point Table for your I Decision presented above. You will be asked to say, on a scale of 0 to 100, how likely you think it is that your Opposite Number will choose Response A.]

(Screen after waiting period for Rounds 2 and 4:)

In this Round's I-Decision you have a new Opposite Number. Your Point Table is unchanged, as shown above. Please look at it. [2 & 3 In a moment you will be asked to say, on a scale of 0 to 100, how likely do you think it is that your Opposite Number will choose Response A.]

(Screen after waiting period for Round 3:)

This Stage is just like the last one except that the I Decision is for a different Mover-Responder problem, Problem 2. Your Point Table for Problem 2 is the one shown above. Please look at it carefully. If you wish to compare it with the one for Problem 1, you can view the latter by clicking the button marked Problem 1. [2 & 3 In a moment, you will be asked how likely do you think it is that your Opposite Number will choose Response A.]

[2 & 3 YOUR L DECISION FOR THIS ROUND. On a scale of 0 to 100, how likely do you think it is that your Opposite Number will choose Response A?

If this Round turns out to be your L-Payment Round, you will be paid for your Report in the way we have explained. It therefore pays you to report accurately.

Try to make a decision within a few minutes. But do not rush, and ask help if you need it. When everyone has decided, we will proceed to your I Decision.

If you need to consult the Table press the 'TABLE' Button

(Subject gives answer).

Please wait for other participants.]

(Screen after waiting period:)

YOUR I DECISION FOR THIS ROUND. Your Opposite Number has now been told the average Report of all the Movers but you, and is now making her/his I Decision for this round. We remind you that the Point Table for the I Decision is the one shown above. Please look at it carefully.

Choose between Move and Pass. Do not rush, but try to make a decision within a few minutes. When everyone has decided, we will proceed to the next Round.

(Subject makes choice).

Please wait for other participants.

(After all participants have finished this round, the process repeats itself from (Loop here after Round 1:) until the end of Round 4)

PLAY STAGE: E PLAYERS

You have been assigned the Responder role.

STAGE 1: GENERAL DESCRIPTION

This Stage has two ROUNDS. In each Round you face the same Mover-Responder problem, Problem 1. In each Round you will make an I Decision with a different Opposite Number. Your Point Table for Problem 1 is shown above.

Your Opposite Number's I Decision is to choose between Move and Pass, and yours is to choose between Response A and Response B. [2 & 3 Before your Opposite Number makes her/his I Decision, she has been asked to make an L Decision.] [3 This is to make a Report on how likely she thinks it is that you will choose Response A.]

[2 & 3 In each Round, before you make your I Decision between Responses A and B, you will be asked to make a G Decision. This G Decision is to make a GUESS about your Opposite Number's L Decision -- that is, to guess her Report on how likely she thinks it is that you will choose Response A. For example, you might think that she has said you are pretty likely to choose Response A -- you might think she gave this a likelihood of 80%; or you might think she has said you are pretty unlikely to -- she gave it a likelihood of 20%. In the former case you would stand to gain most money by answering 80% in your G Decision; in the latter case by answering 20%.]

[3 To help you Guess we are going to tell you what actual people in the Mover role HAVE reported to us. We will not tell you what your own Opposite Number reported, but we will tell you the AVERAGE Report of the other three Movers in this Round. That is, if your Opposite Number reported the likelihood R1 and the other Movers report the numbers R2, R3 and R4, you will be informed of the average of R2, R3 and R4, that is, $(R2 + R3 + R4)/3$.]

To summarize, the procedure for you in each Round is as follows.

- [2 & 3 • Each of the Movers makes an L Decision, giving a Report of how likely she/he thinks it is that you will choose Response A.]
- [3 • You are told the average of the L Decisions of all the Movers except your Opposite Number.]
- [2 & 3 • You make your G Decision, your Guess about the L Decision of your Opposite Number.]

- You make your I Decision between Response A and Response B. At the same time as this, your Opposite Number is making her/his I Decision between Move and Pass.

Until the end of the experiment, nobody will be told anything about anyone else's I Decisions [3, and you will be told nothing about your own Opposite Numbers' L Decisions]. Hence you will learn nothing about outcomes of your I Decisions [2 & 3 or of your G and C Decisions] until the end.

Like Stage 1, Stage 2 consists of two Rounds, in each of which you make [2 & 3 a G Decision and] an I Decision. In each Round you will have an Opposite Number with whom you have not previously interacted in the experiment. The only difference from Stage 1 is in the Mover-Responder problem, which is a different variant, Problem 2, with a different Point Table.

At the end of the experiment your total payment will be determined as follows. The computer programme will randomly choose one of the four rounds of Stages 1 and 2 as your [2 & 3 G-PAYMENT ROUND, and a different round as your] I-PAYMENT ROUND. [2 & 3 You will be reminded of your G Decision in your G-Payment Round. You will then be shown the actual L Decision of your Opposite Number in that round and your resulting G-Payment.] You will be reminded of your I Decision in your I-Payment Round and shown the choice of your Opposite Number in that round and the number of Points you scored. Your final payment will be the sum of

- your Starting Credit of £4,
- [2 & 3 • your G-Payment in your G-Payment Round,]
- the Points you scored in your I Decision in your I-Payment Round, converted into money at £1 per Point.

When everyone is ready, we will begin Round 1 of Stage 1. You may have to wait one or two minutes at this point. We ask you to be patient. Be sure you have understood the whole procedure, referring to Help if you wish to, before clicking on Continue. Put up your hand if you need any further help.

(Loop here after Round 1:)

[2 & 3 *(Screen after waiting period for Round 1:)*

Please look at the Point Table above for this Round's I Decision. You will be asked to make a Guess about what your Opposite Number answered to the question: 'On a scale of 0 to 100, how likely do you think it is that your Opposite Number will choose Response A?']

(Screen after waiting period for Rounds 2 and 4:)

In this Round's I Decision you have a new Opposite Number. Your Point Table is unchanged, as shown above. Please look at it. [2 & 3 In a moment you will be asked to make a Guess about what your Opposite Number answered to the question: 'On a scale of 0 to 100, how likely do you think it is that your Opposite Number will choose Response A?']

(Screen after waiting period for Round 3:)

This Stage is just like the last one except that the I Decision is for a different Mover-Responder problem, Problem 2. Your Point Table for Problem 2 is the one shown above. Please look at it carefully. If you wish to compare it with the one for Problem 1, you can view the latter by clicking the button marked Problem 1. [2 & 3 In a moment, you will be asked to make a Guess about what your Opposite Number answered to the question: 'On a scale of 0 to 100, how likely do you think it is that your Opposite Number will choose Response A?']

Please wait for other participants.

[3 The average report of all Movers other than your current Opposite Number to the question 'On a scale of 0 to 100, how likely do you think it is that your Opposite Number will choose Response A?' is <average>.]

[2 & 3 YOUR G DECISION FOR THIS ROUND. Your Opposite Number has just made an L Decision. It was: 'On a scale of 0 to 100, how likely do you think it is that your Opposite Number will choose Response A?' On a scale from 0 to 100, please make your Guess what her/his Report

was. If this Round turns out to be your G-Payment Round you will be paid for this Guess, and the amount will depend on how accurate it is.

Try to make a decision within a few minutes. But do not rush, and ask help if you need it. When everyone has decided, we will proceed to your I Decision.

If you need to consult the Table press the 'TABLE' Button

(Subject gives answer).]

YOUR I DECISION FOR THIS ROUND. We remind you that the Point Table for the I Decision is the one shown above.

Your Opposite Number is now making her/his I Decision for this round. Please choose between Response A and Response B. Do not rush, but try to make a decision within a few minutes. We will then proceed to the next Round.

(Subject makes choice).

Please wait for other participants.

(After all participants have finished this round, the process repeats itself from (Loop here after Round 1:) until the end of Round 4)

PAYMENT STAGE: R PLAYERS

The Computer chose [2 & 3 Round (*chosen round*) as your L-Payment Round and] Round (*other chosen round*) as your I-Payment Round. The result is as follows:

[2 & 3 In Round (*chosen round*) your Report of the likelihood that your Opposite Number would choose Response A was (*report amount*) and he/she chose RESPONSE (*A or B*). Your L-Payment is therefore (*payment amount*).]

In Round (*other chosen round*) you chose (*Move or Pass*) and your Opposite Number chose RESPONSE (*A or B*). The Problem was Problem (*1 or 2*). Your I-Payment is therefore £ (*payment amount*).

Your Starting Credit was £ 4. Thus your total payment will be £ (*total payment amount*).

Thank you for participating in the experiment. Please wait for an experimenter to come to you.

PAYMENT STAGE: E PLAYERS

You have completed the session. You will learn your payoff when everyone has completed it. Please wait.

The Computer chose [2 & 3 Round (*chosen round*) as your G-Payment Round and] Round (*other chosen round*) as your I-Payment Round. The result is as follows:

[2 & 3 In Round (*chosen round*) your Guess about your Opposite Number's Report of the likelihood you would choose Response A was (*guess amount*) and her/his Report was (*report amount*). Your G-Payment is therefore £ (*payment amount*).]

In Round (*other chosen round*) you chose RESPONSE (*A or B*) and your Opposite Number chose (*Move or Pass*). The Problem was Problem (*1 or 2*). Your I-Payment is therefore £ (*payment amount*).

Your Starting Credit was £4. Thus your total payment will be £ (*total payment amount*).

Thank you for participating in the experiment. Please wait for an experimenter to come to you.

Appendix B. Experimental Data

Trusters						Trustees					
Session	Subject	Round	Game	Statement	Trust	Session	Subject	Round	Game	Guess	Fulfill (F)
1.1	R1	1	K	n/a	0	1.1	E1	1	K	n/a	0
1.1	R1	2	K	n/a	0	1.1	E1	4	N	n/a	0
1.1	R1	3	N	n/a	0	1.1	E1	3	N	n/a	0
1.1	R1	4	N	n/a	0	1.1	E1	2	K	n/a	0
1.1	R2	4	N	n/a	0	1.1	E2	2	K	n/a	0
1.1	R2	1	K	n/a	0	1.1	E2	1	K	n/a	1
1.1	R2	2	K	n/a	0	1.1	E2	4	N	n/a	0
1.1	R2	3	N	n/a	0	1.1	E2	3	N	n/a	1
1.1	R3	3	N	n/a	1	1.1	E3	3	N	n/a	0
1.1	R3	4	N	n/a	1	1.1	E3	2	K	n/a	0
1.1	R3	1	K	n/a	1	1.1	E3	1	K	n/a	0
1.1	R3	2	K	n/a	0	1.1	E3	4	N	n/a	0
1.1	R4	2	K	n/a	1	1.1	E4	4	N	n/a	1
1.1	R4	3	N	n/a	1	1.1	E4	3	N	n/a	1
1.1	R4	4	N	n/a	1	1.1	E4	2	K	n/a	0
1.1	R4	1	K	n/a	1	1.1	E4	1	K	n/a	1
1.2	R5	1	N	n/a	1	1.2	E5	1	N	n/a	0
1.2	R5	2	N	n/a	1	1.2	E5	4	K	n/a	1
1.2	R5	3	K	n/a	1	1.2	E5	3	K	n/a	0
1.2	R5	4	K	n/a	1	1.2	E5	2	N	n/a	0
1.2	R6	4	K	n/a	0	1.2	E6	2	N	n/a	1
1.2	R6	1	N	n/a	1	1.2	E6	1	N	n/a	0
1.2	R6	2	N	n/a	1	1.2	E6	4	K	n/a	1
1.2	R6	3	K	n/a	1	1.2	E6	3	K	n/a	1
1.2	R7	3	K	n/a	1	1.2	E7	3	K	n/a	0
1.2	R7	4	K	n/a	1	1.2	E7	2	N	n/a	0
1.2	R7	1	N	n/a	1	1.2	E7	1	N	n/a	0
1.2	R7	2	N	n/a	1	1.2	E7	4	K	n/a	0
1.2	R8	2	N	n/a	1	1.2	E8	4	K	n/a	1
1.2	R8	3	K	n/a	1	1.2	E8	3	K	n/a	1
1.2	R8	4	K	n/a	1	1.2	E8	2	N	n/a	0
1.2	R8	1	N	n/a	1	1.2	E8	1	N	n/a	1
2.1	R9	1	K	74	1	2.1	E9	1	K	23	0
2.1	R9	2	K	25	1	2.1	E9	4	N	18	0
2.1	R9	3	N	50	1	2.1	E9	3	N	14	1
2.1	R9	4	N	50	1	2.1	E9	2	K	24	0
2.1	R10	4	N	50	1	2.1	E10	2	K	25	1
2.1	R10	1	K	25	1	2.1	E10	1	K	20	1
2.1	R10	2	K	25	0	2.1	E10	4	N	10	0
2.1	R10	3	N	50	1	2.1	E10	3	N	15	0
2.1	R11	3	N	66	1	2.1	E11	3	N	50	0
2.1	R11	4	N	67	1	2.1	E11	2	K	67	1
2.1	R11	1	K	33	0	2.1	E11	1	K	66	1
2.1	R11	2	K	33	0	2.1	E11	4	N	50	0
2.1	R12	2	K	0	0	2.1	E12	4	N	30	1
2.1	R12	3	N	0	0	2.1	E12	3	N	83	1
2.1	R12	4	N	0	0	2.1	E12	2	K	70	1
2.1	R12	1	K	0	0	2.1	E12	1	K	75	1
2.2	R13	1	N	0	0	2.2	E13	1	N	60	1
2.2	R13	2	N	10	0	2.2	E13	2	K	60	0
2.2	R13	3	K	75	1	2.2	E13	3	K	20	0
2.2	R13	4	K	100	0	2.2	E13	4	N	25	0
2.2	R14	4	K	15	0	2.2	E14	1	N	25	0
2.2	R14	1	N	7	0	2.2	E14	2	N	5	0
2.2	R14	2	N	10	0	2.2	E14	3	K	25	1
2.2	R14	3	K	20	0	2.2	E14	4	K	35	1

(continues)

Trusters						Trustees					
Session	Subject	Round	Game	Statement	Trust	Session	Subject	Round	Game	Guess	Fulfill (F)
2.2	R15	3	K	20	1	2.2	E15	1	N	25	0
2.2	R15	4	K	5	0	2.2	E15	2	K	10	0
2.2	R15	1	N	86	1	2.2	E15	3	K	20	0
2.2	R15	2	N	15	1	2.2	E15	4	N	25	1
2.2	R16	2	N	50	0	2.2	E16	1	N	75	1
2.2	R16	3	K	15	0	2.2	E16	2	N	50	0
2.2	R16	4	K	1	0	2.2	E16	3	K	66	0
2.2	R16	1	N	50	0	2.2	E16	4	K	40	0
2.3	R17	1	K	0	0	2.3	E17	1	N	100	1
2.3	R17	2	K	0	0	2.3	E17	4	K	0	0
2.3	R17	3	N	85	1	2.3	E17	3	K	0	0
2.3	R17	4	N	95	1	2.3	E17	2	N	100	1
2.3	R18	1	K	25	1	2.3	E18	2	N	40	0
2.3	R18	2	N	15	1	2.3	E18	1	N	31	0
2.3	R18	3	N	0	1	2.3	E18	4	K	30	0
2.3	R18	4	K	0	1	2.3	E18	3	K	27	0
2.3	R19	1	K	33	0	2.3	E19	3	K	23	0
2.3	R19	2	N	35	0	2.3	E19	2	N	15	0
2.3	R19	3	N	45	1	2.3	E19	1	N	15	0
2.3	R19	4	K	57	1	2.3	E19	4	K	23	0
2.3	R20	1	K	70	1	2.3	E20	4	K	25	0
2.3	R20	2	K	85	1	2.3	E20	3	K	28	0
2.3	R20	3	N	25	1	2.3	E20	2	N	35	0
2.3	R20	4	N	40	1	2.3	E20	1	N	35	0
2.4	R21	1	N	60	1	2.4	E21	1	K	25	0
2.4	R21	2	N	25	1	2.4	E21	2	K	20	0
2.4	R21	3	K	15	0	2.4	E21	3	N	20	0
2.4	R21	4	K	30	1	2.4	E21	4	N	25	0
2.4	R22	1	K	45	1	2.4	E22	1	N	50	1
2.4	R22	2	N	40	1	2.4	E22	2	K	60	1
2.4	R22	3	N	33	0	2.4	E22	3	K	30	1
2.4	R22	4	K	25	0	2.4	E22	4	N	20	0
2.4	R23	1	N	20	0	2.4	E23	1	N	15	0
2.4	R23	2	N	30	0	2.4	E23	2	K	10	0
2.4	R23	3	K	20	0	2.4	E23	3	K	33	1
2.4	R23	4	K	10	0	2.4	E23	4	N	40	1
2.4	R24	1	K	10	0	2.4	E24	1	K	5	0
2.4	R24	2	N	20	0	2.4	E24	2	K	10	0
2.4	R24	3	N	60	1	2.4	E24	3	N	25	0
2.4	R24	4	K	90	1	2.4	E24	4	N	23	0
3.1	R25	1	K	13	1	3.1	E25	1	K	10	0
3.1	R25	2	K	95	0	3.1	E25	2	K	20	0
3.1	R25	3	N	2	1	3.1	E25	3	N	10	0
3.1	R25	4	N	3	0	3.1	E25	4	N	34	0
3.1	R26	1	K	25	1	3.1	E26	1	K	20	1
3.1	R26	2	K	25	0	3.1	E26	2	K	50	1
3.1	R26	3	N	30	1	3.1	E26	3	N	20	1
3.1	R26	4	N	10	1	3.1	E26	4	N	50	1
3.1	R27	1	K	0	0	3.1	E27	1	K	80	1
3.1	R27	2	K	50	0	3.1	E27	2	K	99	1
3.1	R27	3	N	0	1	3.1	E27	3	N	50	1
3.1	R27	4	N	90	0	3.1	E27	4	N	75	1
3.1	R28	1	K	5	0	3.1	E28	1	K	10	0
3.1	R28	2	K	5	0	3.1	E28	2	K	30	1
3.1	R28	3	N	5	0	3.1	E28	3	N	13	1
3.1	R28	4	N	5	1	3.1	E28	4	N	8	1

(continues)

Trusters						Trustees					
Session	Subject	Round	Game	Statement	Trust	Session	Subject	Round	Game	Guess	Fulfill (F)
3.2	R29	1	N	1	1	3.2	E29	1	N	38	1
3.2	R29	2	N	20	1	3.2	E29	2	N	38	1
3.2	R29	3	K	75	1	3.2	E29	3	K	46	0
3.2	R29	4	K	60	1	3.2	E29	4	K	38	0
3.2	R30	1	N	25	1	3.2	E30	1	N	10	0
3.2	R30	2	N	10	0	3.2	E30	2	N	20	0
3.2	R30	3	K	90	1	3.2	E30	3	K	33	0
3.2	R30	4	K	50	1	3.2	E30	4	K	50	0
3.2	R31	1	N	15	1	3.2	E31	1	N	10	0
3.2	R31	2	N	35	1	3.2	E31	2	N	25	0
3.2	R31	3	K	65	1	3.2	E31	3	K	40	0
3.2	R31	4	K	55	1	3.2	E31	4	K	40	0
3.2	R32	1	N	20	0	3.2	E32	1	N	20	0
3.2	R32	2	N	10	0	3.2	E32	2	N	20	1
3.2	R32	3	K	12	0	3.2	E32	3	K	70	1
3.2	R32	4	K	8	0	3.2	E32	4	K	75	1
3.3	R33	1	N	65	1	3.3	E33	1	N	93	1
3.3	R33	2	N	65	1	3.3	E33	2	N	97	1
3.3	R33	3	K	65	1	3.3	E33	3	K	95	1
3.3	R33	4	K	65	1	3.3	E33	4	K	96	1
3.3	R34	1	N	90	0	3.3	E34	1	N	25	0
3.3	R34	2	N	90	1	3.3	E34	2	N	60	1
3.3	R34	3	K	80	1	3.3	E34	3	K	50	0
3.3	R34	4	K	40	1	3.3	E34	4	K	50	0
3.3	R35	1	N	33	1	3.3	E35	1	N	75	0
3.3	R35	2	N	25	0	3.3	E35	2	N	25	0
3.3	R35	3	K	15	0	3.3	E35	3	K	65	0
3.3	R35	4	K	33	1	3.3	E35	4	K	65	0
3.3	R36	1	N	50	1	3.3	E36	1	N	66	0
3.3	R36	2	N	25	1	3.3	E36	2	N	70	0
3.3	R36	3	K	50	0	3.3	E36	3	K	32	0
3.3	R36	4	K	29	1	3.3	E36	4	K	35	0
3.4	R37	1	K	0	1	3.4	E37	1	K	55	1
3.4	R37	2	K	88	1	3.4	E37	2	K	66	1
3.4	R37	3	N	80	1	3.4	E37	3	N	60	1
3.4	R37	4	N	100	1	3.4	E37	4	N	66	1
3.4	R38	1	K	20	0	3.4	E38	1	K	30	0
3.4	R38	2	K	20	0	3.4	E38	2	K	80	1
3.4	R38	3	N	0	0	3.4	E38	3	N	90	1
3.4	R38	4	N	0	0	3.4	E38	4	N	75	1
3.4	R39	1	K	75	1	3.4	E39	1	K	75	1
3.4	R39	2	K	90	1	3.4	E39	2	K	68	0
3.4	R39	3	N	100	1	3.4	E39	3	N	70	1
3.4	R39	4	N	50	1	3.4	E39	4	N	33	1
						3.4	E40	1	K	25	0
						3.4	E40	2	K	100	0
						3.4	E40	3	N	25	0
						3.4	E40	4	N	35	0

The session numbers are in the format $p.q$, where p is the condition number (1, 2 or 3) and q is an indicator for the specific session. The condition 3 sessions are the same as in Bacharach et al. (2001, p. 18), namely their sessions 2, 6, 8 and 10 (in their session 10, or session 3.4 here, a trustor's computer failed, and so there is data for one subject less). K and N are the game codes for the KTG and NTG, respectively. Statements and guesses are unavailable (n/a) for condition 1 sessions, since these sessions had no belief elicitation. *Trust* is equal to 1 if trustor placed trust, 0 otherwise; *Fulfill (F)* is equal to 1 if the trustee fulfills trust, 0 otherwise.

Fig. 1. The Basic Trust Game (BTG): general case and two specific variants.

(a) Basic Trust Game (BTG)

$$y < a, a < w, x < z$$

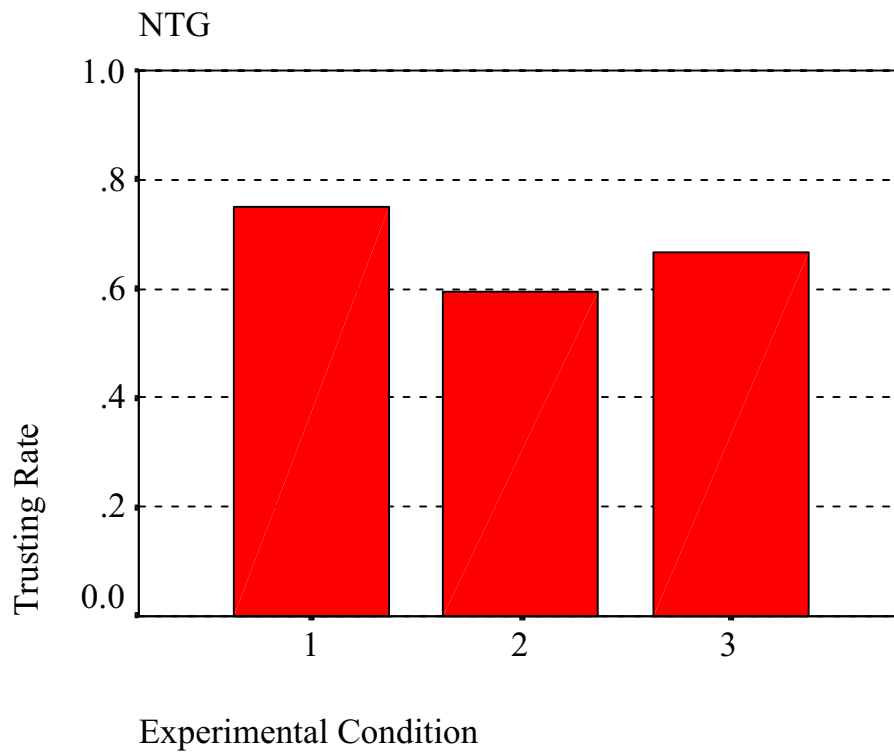
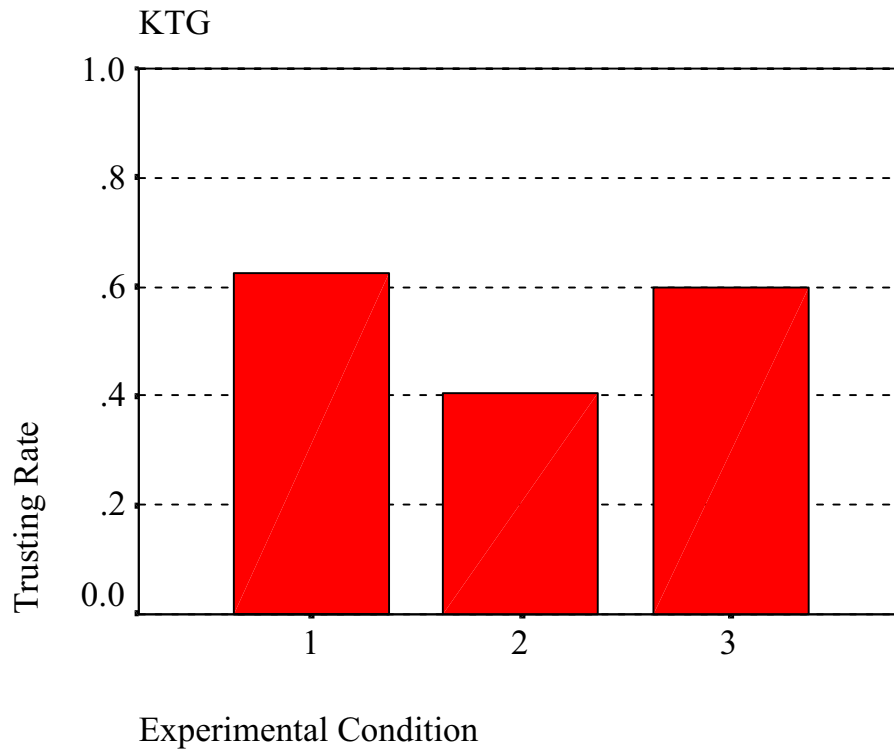
		Trustee	
		<i>Fulfill (F)</i>	<i>Violate</i>
Truster	<i>Trust</i>	w, x	y, z
	<i>Withhold</i>	a, b	a, b

(b) BTG Variant 1: Kind Trust Game (KTG)

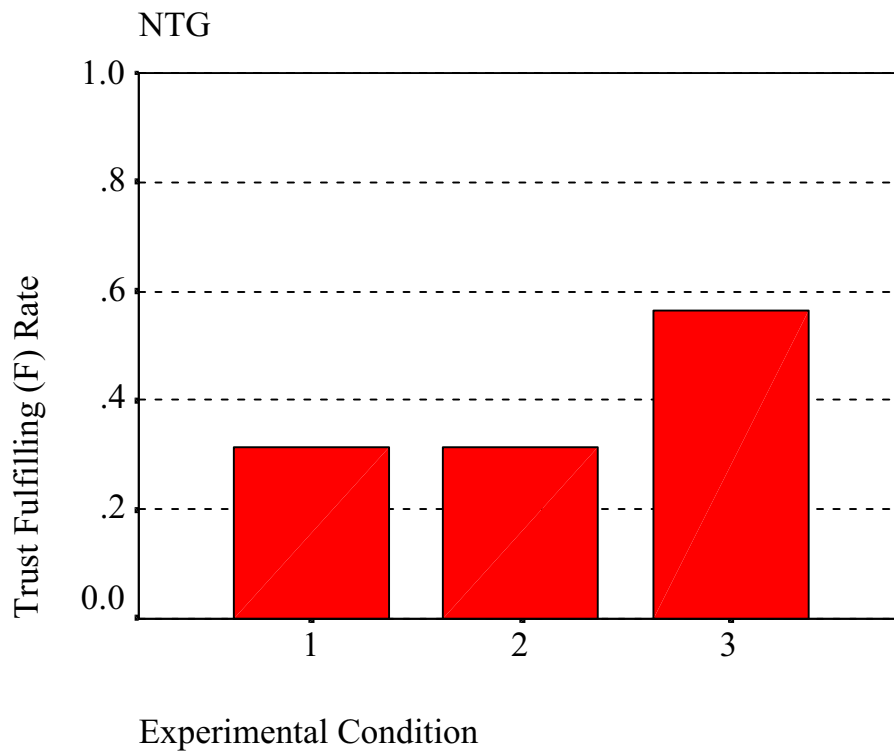
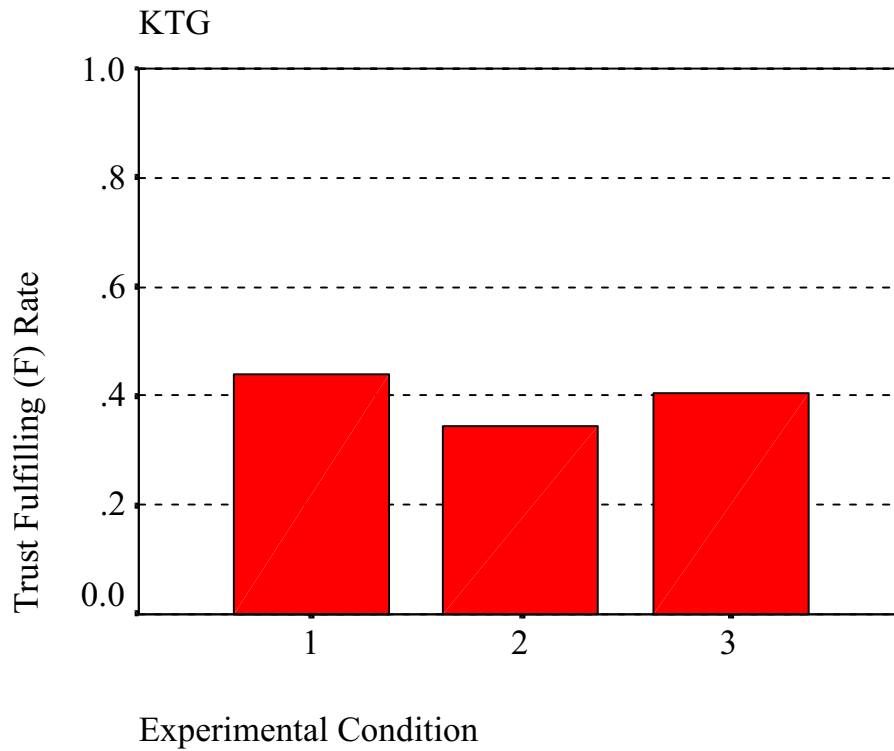
		Trustee	
		<i>Fulfill (F)</i>	<i>Violate</i>
Truster	<i>Trust</i>	2, 2	0, 3
	<i>Withhold</i>	1, 1	1, 1

(c) BTG Variant 2: Needy Trust Game (NTG)

		Trustee	
		<i>Fulfill (F)</i>	<i>Violate</i>
Truster	<i>Trust</i>	3, 3	0, 4
	<i>Withhold</i>	1, 2	1, 2

Fig. 2. Trusting rates by trusters in the KTG and the NTG.

Condition 1: belief elicitation and transmission. Condition 2: belief elicitation only. Condition 3: no belief elicitation.

Fig. 3. Trust fulfilling rates by trustees in the KTG and the NTG.

Condition 1: no belief elicitation. Condition 2: belief elicitation. Condition 3: belief elicitation *and* transmission.

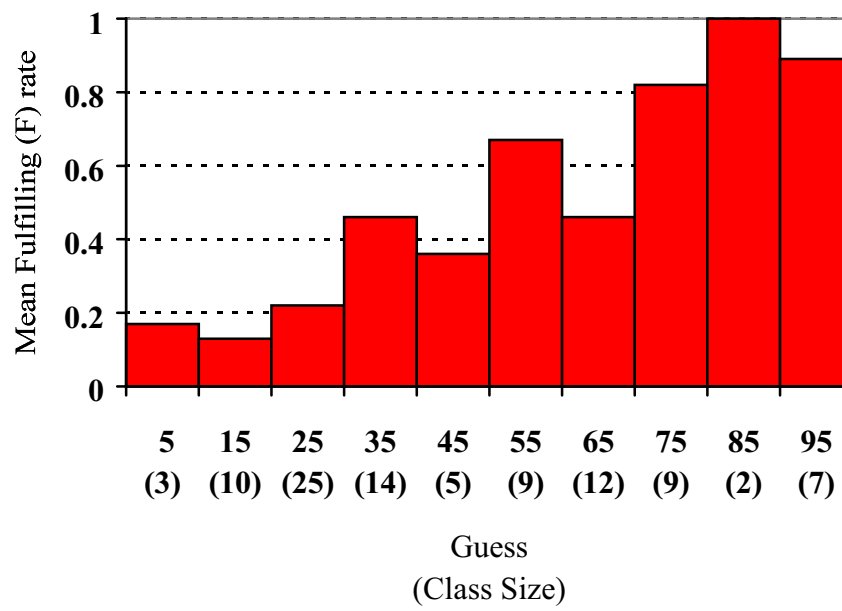
Fig. 4. Variation of fulfilling rate with guess.

Table 1
Observed trusting and fulfilling rates by game and condition

Condition	1	2	3
KTG Trusting Rate	0.625	0.406	0.600
NTG Trusting Rate	0.750	0.594	0.667
Aggregate Trusting Rate	0.688	0.500	0.633
KTG Fulfilling Rate	0.438	0.344	0.406
NTG Fulfilling Rate	0.313	0.313	0.563
Aggregate Fulfilling Rate	0.375	0.328	0.484

Table 2
Probit regression analysis on F (the choice to fulfill)

Explanatory Variables	Model 1				Model 2			
	Coeff.	S.E.	Prob.	Sign.	Coeff.	S.E.	Prob.	Sign.
Ntg	0.36	0.544	0.508					
Communication	0.668	0.596	0.262					
Ktg x Guess	0.045	0.013	0.001	*	0.025	0.006	0	*
Ntg x Guess	0.033	0.008	0	*	0.031	0.007	0	*
Communication x Guess	-0.024	0.014	0.078					
Communication x Ntg x Guess	0.016	0.012	0.183					
Constant	-1.993	0.482	0	*	-1.401	0.302	0	*

These probit regressions are on a dummy, F , equal to 1 when the trustee chooses to fulfill, and to 0 otherwise ($n = 128$). The log-likelihood is equal to -67.226 and -69.370 for model 1 and 2, respectively. Ntg = 1 when the NTG is played (0 otherwise). Ktg = 1 when the KTG is played (0 otherwise). Communication = 1 in condition 3 (0 in condition 2). Guess is the revealed guess f^{**} by the trustee. Standard errors are corrected for individual-specific effects. * stands for significance at the 0.01 level.