Testing Guilt Aversion

Tore Ellingsen,
Department of Economics, Stockholm School of Economics,
Box 6501, SE-113 83 Stockholm, Sweden.
tore.ellingsen@hhs.se

Magnus Johannesson,
Department of Economics, Stockholm School of Economics,
Box 6501, SE-113 83 Stockholm, Sweden.
magnus.johannesson@hhs.se

Sigve Tjøtta,
Department of Economics, University of Bergen,
Hermann Fossgate 6, 5007 Bergen, Norway.
sigve.tjotta@econ.uib.no

Gaute Torsvik,
Department of Economics, University of Bergen,
Hermann Fossgate 6, 5007 Bergen, Norway.
gaute.torsvik@econ.uib.no

SSE/EFI Working Paper Series in Economics and Finance
No. 683
December, 2007

Abstract

Guilt averse individuals experience a utility loss if they believe they let someone down. In particular, generosity depends on what the donor believes that the recipient expects to receive. In experimental work, several authors have identified a positive correlation between such second-order donor beliefs and generous behavior, as predicted by the guilt aversion hypothesis. However, the correlation could alternatively be due to a "false consensus effect," i.e., the tendency of people to believe others to think like themselves. In order to test the guilt aversion hypothesis more rigorously, we conduct three separate experiments: a dictator game experiment, a complete information trust game experiment, and a hidden action trust game experiment. In the experiments we inform donors about the beliefs of their respective recipients, while eliciting these beliefs so as to maximize recipient honesty. The correlation between generous behavior and donors' second-order beliefs is close to zero in all three experiments.

Key words: guilt aversion, beliefs, generosity, experiments.
JEl Codes: C91, D64.

Acknowledgement: Thanks to Gary Charness, Martin Dufwenberg, and Robert Östling for helpful comments on earlier drafts and to Anna Breman, Stein Svalestad and Niklas Zethraeus for research assistance. Financial support from The Jan Wallander and Tom Hedelius Foundation, the Swedish Research Council, and the Norwegian Research Council is gratefully acknowledged.

1. INTRODUCTION

How should economists model generous and trustworthy behavior? A growing body of evidence suggests that generosity and trustworthiness are not simply features of people's preferences over consumption allocations, or even over transfers.[1] Social context matters as well. Specifically, people are concerned about others' *beliefs*. "What is expected of me?" "What will others think?"

Two lines of theory formalize the intuition that people get utility from their beliefs about others' beliefs, that is, from second-order beliefs. In the *social esteem model*, people care about what others think about them – feeling proud if others think highly of them and shameful otherwise.[2] In the *guilt aversion model*, people care about what others expect of them, feeling guilty if their behavior falls short of expectations; see Charness and Dufwenberg (2006) for an intuitive account and Battigalli and Dufwenberg (2005, 2007) for formalities.[3]

The social esteem model is relatively easy to test by varying actor anonymity and/or observability of actions.[4] Guilt aversion is harder to identify empirically. Guilt depends on one's beliefs about others' anticipations, and both anticipations and beliefs about anticipations are difficult for the researcher to manipulate. Therefore, the existence of guilt aversion has usually been inferred by asking subjects what they believe that their opponents expect; see Bacharach, Guerra and Zizzo (2007), Charness and Dufwenberg (2006), Dufwenberg and

---

[1] Pure altruism, as analyzed by Becker (1974), entails preferences over allocations only. Impure altruism, as analyzed by Andreoni (1989,1990), also admits preferences over transfers.

[2] Following Bernheim (1994), Ireland (1994), and Glazer and Konrad (1996), economists have used signaling models to study concerns for social esteem. Recent theoretical contributions include Andreoni and Bernheim (2007), Benabou and Tirole (2006), Ellingsen and Johannesson (forthcoming), and Tadelis (2007).

[3] Battigalli and Dufwenberg (2005,2007) build on the "psychological game theory" framework pioneered by Geanakoplos, Pearce, and Stacchetti (1989) (GPS) and Rabin (1993). Interestingly, in their intuitive justification for their theory GPS emphasize social esteem (pride and shame), not guilt: a player gets utility from "what he thinks his friends will think about his character" (page 66). For extensive surveys of social preference theories, including more general discussions of the role second-order beliefs, we refer to Sobel (2005) and Fehr and Schmidt (2006).

[4] For relevant experimental evidence, see for example Andreoni and Bernheim (2007), Cox and Deck (2005), Dana, Cain, and Dawes (2006), Hoffman, McCabe and Smith (1996), and Tadelis (2007).

Gneezy (2000), and Guerra and Zizzo (2004).[5] The studies reveal significant correlations between second-order beliefs and actions. The evidence is clearly consistent with the guilt aversion hypothesis. However, the test may be too weak. As Charness and Dufwenberg (2006, page 1594) note in the discussion of their trust game evidence, an alternative explanation is that trustees believe that other trustees would choose like them, and that trustors' beliefs lean in this direction too – essentially a false consensus effect (Ross, Greene, and House, 1977):[6] Trustees who prefer to make larger back-transfers to the trustor, believe that trustors expect large back-transfers.

In the present paper we attempt to test the guilt aversion hypothesis more rigorously by eliminating such false consensus effects. The tests work by providing information to each player about the paired player's first-order beliefs.[7] Our simplest experiment elicits recipient beliefs in a Dictator game and communicates these beliefs to the dictator before the dictator makes the allocation decision.[8] The guilt aversion hypothesis postulates a positive relationship between the elicited belief and the amount of money allocated to the recipient, and this stronger test does not admit a reverse causality interpretation.

If recipients realize that their beliefs will be communicated to the dictator, reported beliefs may be untruthful. Indeed, if recipients believe that dictators are guilt averse, reports should be distorted upwards, potentially heavily so. In order to eliminate or at least mitigate such strategic reporting, we do not tell recipients that the dictators will have access to their

---

[5] For a similar test in a public goods game, see Dufwenberg, Gächter, and Hennig-Schmidt (2007).
[6] The false consensus effect is the propensity to believe that others' perceptions and thoughts are similar to one's own. When trustees are asked what they think that trustors think, the false consensus effect implies that their assessment is contaminated by their own belief about what other trustees will do. If this belief in turn is correlated with the own planned action – for any reason – we have a reverse causation problem.
[7] Bacharach, Guerra and Zizzo (2007) and Guerra and Zizzo (2004) have previously elicited predictions from trustors. However, they informed trustees only about the mean prediction made by trustors with whom they were not matched. They do not report correlations between trustee behavior and this average.
[8] Dufwenberg and Gneezy (2000) also used a Dictator game to investigate guilt aversion. However, they did not elicit beliefs from recipients – only second-order beliefs from dictators.

beliefs. We merely ask them to guess the outcome (and even provide a material incentive to guess accurately).

Our dictator game evidence reveals virtually zero correlation between the recipient's guess and the dictator's allocation decision. Since the Dictator game is quite extreme, it would nevertheless be premature to reject the guilt aversion hypothesis on the basis of this experiment alone. Therefore we also conduct two different trust game experiments, the latter of which is virtually identical with the hidden action trust game considered by Charness and Dufwenberg (2006). In these games, we again elicit trustors' expectations and investigate to what extent they affect trustees' behavior. As in the Dictator game, the correlation is close to zero. We conclude that guilt aversion appears to be an unimportant factor in these experiments, and that previous findings to the contrary may be driven by false consensus effects.

2. EXPERIMENT I

The first experiment is a double-blind dictator game, in which neither other subjects nor the experimenters can observe the decision of a particular subject (Hoffman et al., 1994,1996).[9]

2.1 Design

One subject (the dictator) decides how to allocate SEK 120 between herself and another subject (the recipient) in another room (SEK=Swedish Kronor; $1≈ SEK 8 at the time of the experiment).

[9] A difference compared to previous double-blind dictator game experiments is that no dictator received an envelope without any money (this procedure has previously been used as an extra guarantee of anonymity, as

To identify the effect of an exogenous variation in second-order beliefs, we elicit recipients' beliefs about donations prior to the dictator game. Every recipient guesses how much dictators will give on average, and the best guess is rewarded a SEK 100 payment after the experiment.[10] Prior to making the allocation decision, the dictator is shown the guess of their recipient.

To counter the risk that recipients' guesses are contaminated by a desire to influence the dictator, recipients are not told that their guess will be shown to the dictator.[11] If second-order beliefs are important for donations as predicted by guilt aversion, there should be a positive correlation between the dictator allocation and the recipient's guess.[12]

The subjects were recently enrolled undergraduate students at the Stockholm School of Economics in Sweden. We conducted three sessions of the experiment and a total of 171 subjects participated. Three of these were used as monitors (see below). The remaining 168 subjects yielded 84 pairs of observations. The experimental procedures are further described below (the complete instructions are reproduced in Appendix I).

---

even if no dictator donates any money the experimenter will not be able to infer the decision of a specific subject).

[10] In principle, the optimal guess may differ slightly from one's expectation of the average, since the probability of winning depends on the distribution of other guesses. However, even if subjects would have identical expectations the theoretical effect is minor, and we see no evidence suggesting that it is important in practice; for example, all subjects except one make integer guesses, which they should not if they worried about ties. At any rate our test primarily requires that the guess is positively correlated with the true belief. (Note also that our other two experiments avoid the whole issue by rewarding guesses based on absolute rather than relative performance.)

[11] The dictator is also informed that the recipient made the guess without knowing that it would be shown to the dictator.

[12] Anchoring provides an alternative hypothesis for why we would observe a positive correlation in the experiment. Research in psychology suggests that subjects may anchor on numbers or monetary amounts even if they provide no valuable information (Tversky and Kahneman, 1974; Ariely, Loewenstein, and Prelec, 2003). To assess the size of any anchoring effects we therefore also conducted a control treatment (also with separate subjects), in which the subjects participated in a guessing game prior to the dictator game. They were asked to guess a randomly drawn amount between SEK 0 and SEK 120, and the closest guess was rewarded a SEK 100 payment. These guesses were shown to the dictators before they made their allocation decisions. If there is an anchoring effect, the dictator donations will be correlated with these guesses. We found no evidence of anchoring.

2.2 Procedures

Subjects are recruited to two separate rooms called room A and room B. Dictators are in room A and recipients are in room B. The subjects are welcomed and told not to talk to each other. In room B (recipients), subjects receive instructions for eliciting beliefs (i.e. guessing the outcome of the dictator game they will subsequently play). In room A (dictators), subjects receive the dictator game instructions. The subjects read the instructions. Thereafter they can ask questions individually.

In room B, subjects enter their guess of the mean allocation and their student identification number on a numbered form marked "guess". The experimenter collects the forms and removes the part with the student identification number.[13] Thereafter the experimenter puts each form into an unmarked envelope that contains six SEK 20 bills (and subjects in room B now also receive the instructions for the dictator game experiment). The experimenter brings the envelopes into room A. In room A, a monitor has been chosen among the subjects and he/she conducts the experiment and verifies that the procedures are followed as described in the instructions.[14] The monitor gives one envelope to each subject in room A. Subjects are asked to open the envelope and read the guess on the form and thereafter put the form back into the envelope. Thereafter the monitor calls one person at a time and the subject goes behind a screen. In private behind the screen, the subject decides how many SEK 20 bills to leave in the envelope and how many to keep for his/her own use. The subject then seals the envelope and drops it in a box marked "Mail".

When all subjects in room A have made their decisions, the monitor brings the box marked "Mail" to room B. The monitor opens an envelope to check the number on the form

[13] We need the student identification number to be able to pay the best guess the SEK 100 prize; after the experiment we can match each guess with a specific student identification number by the number on the form (that we write also on the part with the student identification number that is separated).

marked “guess” and ask the person with this number to follow the monitor to an adjacent room. The monitor records the content of the envelope (the sum of money, if any, that is donated) and gives the contents to the person called. That person then leaves the experiment. The monitor continues until all envelopes have been opened. The experiment is then over.

2.3 Results

The distribution of donations is shown in Figure 1. The average donation is 24% of the endowment. Almost two thirds (65%) of the subjects donate something, and nearly 20% of the subjects split the endowment 50/50. Perhaps surprisingly, four subjects (5% of the subjects) donate the entire amount. [insert Figure 1 about here]

The distribution of beliefs is shown in Figure 2. The most common beliefs are 17% (SEK 20) and 33% (SEK 40), followed by 50% (SEK 60). On average the recipients expect to get 32% of the endowment, which is significantly higher than the mean donation of 24% ($p=0.015$ according to an independent samples t-test and $p=0.002$ according to a Mann-Whitney test). In Figure 3 we plot the relationship between donations and beliefs. It is difficult to spot any clear-cut pattern between recipient beliefs and actual donations. The lack of connection is confirmed by the correlation coefficients. The parametric Pearson correlation coefficient is -0.075 and highly non-significant ($p=0.497$); the non-parametric Spearman correlation coefficient is -0.044 ($p=0.689$). As a sensitivity analysis we removed the six observations with beliefs exceeding 50%. The correlation coefficients approach zero (-0.008 (Pearson) and -0.007 (Spearman)). We also tried removing the four subjects that donated everything, but also in this case the correlation coefficients are close to zero (0.012 (Pearson)

---

[14] The monitor receives SEK 120 in payment.

and 0.012 (Spearman)).[15] Finally, we tried removing all 29 subjects that donated nothing – and for whom guilt aversion may be a secondary concern. The Pearson correlation coefficient is then -0.026 (p=0.852) and the Spearman is 0.091 (p=0.509). We therefore cannot reject the null hypothesis of a zero correlation, and fail to find support for guilt aversion. [insert Figure 2 and 3 about here]

3. EXPERIMENT II

The second experiment investigates the effect of second-order beliefs in a double-blind trust game.

Trust games capture the essence of many bilateral economic transactions.[16] One party, the trustor, can take an action that yields an acceptable private return if and only if the other party, the trustee, subsequently refrains from maximizing the own gains. The trustor is a metaphor both for the employee who decides whether or not to invest in relationship-specific human capital and for the employer who decides whether or not to engage in close monitoring of employees.

Trust game behavior is challenging to explain, because the behavior of trustees frequently depends on factors beyond the trustees' choice set. For example, McCabe, Rigdon, and Smith (2003) find that trustees are substantially more likely to reward trust if trustors' outside (no-trust) options are relatively attractive.[17] Such trustee behavior is difficult to reconcile not only with standard selfish preferences, but also with most models of social

---

[15] If we remove both the four subjects that donated everything and the six subjects that received beliefs over 50%, the correlation coefficients are positive (0.099 (Pearson) and 0.051 (Spearman)), but still highly non-significant (p=0.400 and p=0.668, respectively).

[16] Like Camerer (2003) we define trust games quite broadly, including the both the classical investment game of Berg, Dickhaut, and McCabe (1995) and the various simplifications of this game in which each player has fewer (but at least two) actions to choose between.

[17] Related findings are reported by Charness (2004), Cox (2004), Falk and Kosfeld (2006), and Falk, Fehr, and Fischbacher (2007). For a general approach to the problem of disentangling reciprocity from altruism, see Cox, Friedman, and Sadiraj (forthcoming).

preferences. Plain altruism or inequality aversion cannot explain why unchosen options at the initial stages of a game should be relevant for play at later stages. A (different) form of reciprocity appears to be at play. Could it be guilt aversion? Battigalli and Dufwenberg (2005) notes that a person who trusts despite an attractive outside option reveals optimistic expectations, and will be correspondingly disappointed if the trustee fails to reward the trusting action. Trustees who are susceptible to feeling guilty may therefore reward trust in order to avoid the unpleasant sensation of letting the trustor down.

3.1 Design

At the first stage of the game one subject, the trustor, decides how much of an initial endowment of NOK 50 to send to a second player, the trustee (NOK=Norwegian Kroner; $1≈ NOK 6 at the time of the experiment). The sum sent is multiplied by 5. In the second stage, the trustee decides how much of the total sum to send back to the trustor (the back-transfer), but is restricted to choose an amount from the set {0, 50, 100, 150, 200, 250}. Note that the second stage is essentially a dictator game.[18]

To identify the effect of an exogenous variation in second-order beliefs we elicit trustors' beliefs about back-transfers after the trustors have made their investment decision. Every trustor guesses the most common amount a trustee will back-transfer conditional on investment (i.e., they guess the mode of the distribution). A correct guess is rewarded with NOK 50. The trustee is shown the guess of the trustor prior to making the allocation decision. To avoid strategic guesses, we did not inform the trustors that the trustees would observe their guesses. If second-order beliefs are important for back-transfers as predicted by guilt

[18] The experiment is similar to the f4 experiment in Dufwenberg and Gneezy (2000). However, note that the trustee makes the back-transfer decision after learning the investment decision. Dufwenberg and Gneezy instead use the strategy method.

aversion, there should be a positive correlation between the trustor belief and the back-transfer of the trustee.[19]

The subjects were students at the University of Bergen in Norway. We conducted four sessions. A total of 196 subjects participated in the experiment, yielding 98 subject pairs. As we are interested in the behaviour of the trustees, we lose observations for the ten pairs where no investments were made, leaving us with 88 observations in the analysis. The experimental procedures are further described below (the complete instructions are reproduced in Appendix II).

3.2 Procedures

Subjects are welcomed and told no to talk to each other. They are randomly assigned to room A (trustors) and room B (trustees), where they receive experimental instructions. The instructions in room A describe the game, but contain no information about elicitation of beliefs. After the instructions are read and questions answered, the trustors receive two envelopes. One contains a NOK 50 bill and a green piece of paper of the same shape and colour as the NOK 50 bill. The other envelope has an identification number. The trustors then decide whether to place the NOK 50 bill or the green piece of paper in the envelope with the identification number or in a personal envelope. After everyone has made their decision the envelopes are collected in a closed box by the experimenter in room A. The box is then given to an assistant waiting outside the room, who takes the box to another room (C).

The trustors are then asked what they believe is the most common amount that will be paid back to trustors that gave NOK 50 to the trustees. They receive a sheet of paper, marked

---

[19] As for Experiment I, we include a control treatment to control for any anchoring effects. In the control treatment trustors participate in a guessing game prior to the trust game where they are asked to guess on one of the following randomly drawn amounts: NOK 0, NOK 50, NOK 100, NOK 150, NOK 200, NOK 250. Correct guesses are rewarded a NOK 50 payment. These guesses are shown to the trustees before they make their back-

with their identification number, and are instructed to circle the amount they believe will be paid back by most trustees. Subsequently, the guesses are collected by the experimenter in room A, who gives them to the assistant waiting outside the room. In room C, the assistant records investment decisions and guesses and places the guesses in the envelope with a matching identification number. The assistant then brings the box with envelopes to the experimenter in room B.

The experimenter in room B gives one envelope to each subject (trustee) in room B. The envelope either contains NOK 250 and a guess, or a green piece of paper (the fake money) and a guess. The trustees who receive money decide the amount they will return to the trustors and place it in the envelope with the identification number. After the allocation is made, all envelopes with identification numbers are collected in a box by the experimenter in room B. The box is given to the assistant waiting outside the room, who registers the decisions in room C. Finally, the assistant gives the box with the envelopes to the experimenter in room A, who distributes the envelopes to the trustors. The experiment is then over.

3.3 Results

The distribution of back-transfers is shown in Figure 4. The average back-transfer is 43% of the endowment. The most common back-transfers are 40% (NOK 100) or 60% (NOK 150) of the endowment. [insert Figure 4 about here]

The distribution of beliefs in the beliefs group is shown in Figure 5. The most common beliefs are that the trustee will transfer 20%, 40%, or 60%. The average belief is 41%, which

---

transfer decision. If there is an anchoring effect the back-transfers will be correlated with these guesses. We found no evidence of an anchoring effect.

is not significantly different from the mean back-transfer (p=0.287 according to an independent samples t-test and p=0.194 according to a Mann-Whitney test).

In Figure 6 we plot the relationship between back-transfers and beliefs. There is no clear-cut pattern between beliefs and back-transfers in the Figure. The parametric Pearson correlation coefficient is 0.085 (p=0.434) and the non-parametric Spearman correlation coefficient is 0.124 (p=0.249). Removing the subject that transferred everything has little effect on the results (Pearson correlation=0.044 (p=0.683) and Spearman correlation=0.096 (p=0.378)).[20] Removing the two subjects with beliefs of a 100% back-transfer also yields similar results (Pearson correlation=0.122 (p=0.264) and Spearman correlation=0.140 (p=0.197)).[21] We therefore cannot reject the null hypothesis of a zero correlation. [insert Figure 5 and 6 about here]

4. EXPERIMENT III

We finally investigate the presence of guilt aversion in a trust game with hidden actions, as developed by Charness and Dufwenberg (2006). For ease of comparison, we describe the game using the original dollar payoffs.

4.1 Design

There are two players (A and B). In the first stage, player A decides between In and Out. If Out is chosen the game ends and both players receive $5. If In is chosen, player B decides between “Roll” and “Don’t Roll”. If “Don’t Roll” is chosen player A receives $14 and player B receives nothing. If “Roll” is chosen player B always receives $10, but rolls a six-

[20] If we also remove the two subjects that transferred 80%, the Pearson correlation is 0.086 (p=0.432) and the Spearman correlation is 0.138 (p=0.209).

[21] If we remove both the subject transferring 100% and the two subjects with beliefs of 100% back-transfers, the Pearson correlation coefficient is 0.072 (p=0.512) and the Spearman correlation coefficient is 0.109 (p=0.319).

sided die to determine the payoff for player A (with probability 1/6 A receives nothing and with probability 5/6 A receives $12). We use the same relative payoffs, but multiply by 20 to get the payoffs in NOK.

To measure guilt aversion, Charness and Dufwenberg let subject A players guess what fraction of B players will chose Roll. To measure second-order beliefs they then let subject B players guess the beliefs of A, conditional on A selecting to move In. To get an exogenous variation in beliefs we modify this design so that B players instead are informed about the guess (belief) of their paired A player before deciding whether to Roll or not. As Charness and Dufwenberg, we use the strategy method, so that all B players choose between Roll and Don't Roll before they are informed about whether their counterpart chose In or Out. In our experiment Player A first choose In or Out. Player A then guesses what fraction of B players will chose Roll (without being informed that this guess will be given to player B). If the guess is within 5 percentage units of the actual fraction, Player A earns NOK 50. Player B is informed about the guess of player A, and thereafter decides whether to choose Roll or Don't Roll.[22] Finally, Player B is informed about whether player A chose In or Out. According to the guilt aversion hypothesis, there should be a positive correlation between chosing Roll and the second-order belief (the guess of player A).[23]

The subjects were students at the University of Bergen in Norway**.** A total of 88 subjects participated in the experiment, yielding 44 subject pairs. The sample size is thus similar to the sample size in Charness and Dufwenberg, which had 45 observations in the corresponding treatment. The experimental procedures are further described below (the complete instructions are reproduced in Appendix III).

---

[22] Player B is also informed that player A did not know that the guess would be shown to Player B.
[23] Observe that a trustee receives the trustor's guess before knowing whether the trustor has chosen In. This is necessary in order to utilize the strategy method for the trustees, as do Charness and Dufwenberg.

4.2 Procedures

Subjects are welcomed and told no to talk to each other. Subjects are randomly assigned to room A (principals) and room B (agents), and thereafter receive experimental instructions. The instruction in room A describes the game, but contains no information about elicitation of beliefs. The subjects read the instructions and can ask questions individually. After the instructions are read and questions answered, subjects in room A receive a sheet of paper with an identification number. They mark the sheet either IN or OUT. The sheets are collected by the experimenter in room A and handed over to an assistant waiting in room C.

Subjects in room A then receive new instructions and a form with their identification number. They are told to guess the percentage of subjects in room B that will roll the die. They are informed that if their guess is within a margin of 5 percentage units of the true percentage of rolls they will receive 100 NOK. The choices are collected by the experimenter in room A and handed over to the assistant waiting outside the room. The assistant gives the form with the guesses to the experimenter in room B.

After reading the instructions, the subjects in room B receive two forms, one is the form with guesses from room A on which they are informed about the beliefs of their counterpart. On the other form, marked with their identification number, they indicate whether they will ROLL THE DIE or NOT ROLL THE DIE. They are told to fold the form that contains their decision. The forms are then collected by the experimenter in room B. Everyone in room B then rolls the die and the experimenter writes down the outcome. The die is rolled by everyone in order not to reveal the individual choices made in the game to other subjects. The decision forms and the outcome of the die rolls are then given to the assistant waiting outside the room. In room C, the assistant calculates the payoff of each person, and the money is placed in envelopes which the assistant hands over to the experimenters in room A and B. The

experimenters distribute the envelopes to the subjects and tell them that the experiment is over.

4.3 Results

Figure 7 illustrates the decisions in the experiment. At the first stage, 59% of A's choose In. In the second stage, 61% of B's choose Roll. The combination of In and Roll occur in 34% of the subject pairs. These fractions are quite similar to those of Charness and Dufwenberg. [insert Figure 7 about here]

Figure 8 displays the distribution of beliefs. The average belief is 43%, which is lower than the actual rate of Roll.[24] In Figure 9 we plot the relationship between beliefs and the decision to Roll. There is no discernable pattern. The Pearson correlation coefficient is -0.071 (p=0.646) and the Spearman correlation coefficient is -0.066 (p=0.669). Removing the highest and lowest beliefs (95% and 0%) entails similar results; the Pearson correlation coefficient is -0.162 (p=0.306) and the Spearman correlation coefficient is -0.160 (p=0.311). Charness and Dufwenberg did not report correlation coefficients between beliefs and the decision to Roll, but compared the difference in beliefs between those who chose to Roll and those who chose Don't Roll. For the treatment that is closest to ours (no communication and 5,5 as Out payoffs), they reported mean beliefs of 54.2% for Roll and 39.6% for Don't Roll. For the same comparison, we obtain mean beliefs of 41.6% for Roll and 45.7 for Don't Roll. These numbers are not significantly different (p=0.657 according to an independent samples t-test and p=0.664 according to a Mann-Whitney test). The measure of guilt aversion reported by Charness and Dufwenberg thus seem to be affected by the false consensus effect. [insert Figure 8 and 9 about here]

5. CONCLUSION

When decision makers have information about opponents' beliefs, our measure of guilt aversion is insignificantly different from zero in each of our three experiments. The experiments are relatively large, involving a total of 455 subjects in two countries. If Scandinavian students are guilt averse, the effect is small. Generosity and positive reciprocity among these subjects could be caused by preferences for redistribution or by pride and shame, but appears unaffected by guilt from letting the opponent down.

---

[24] This difference is significant at the 5% level with an independent samples t-test ($p=0.038$) and at the 10% level with a Mann-Whitney test ($p=0.052$).

REFERENCES

Andreoni, J. (1989). Giving with impure altruism: Applications to charity and Ricardian equivalence. *Journal of Political Economy* 97, 1447-1458.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm–glow giving. *Economic Journal* 100, 464-477.

Andreoni, J. and Bernheim, B.D. (2007). Social image and the 50:50 norm. Manuscript. UC San Diego.

Ariely, D., Loewenstein, G. and Prelec, D. (2003). Coherent arbitrariness: Stable demand curves without stable preferences. *Quarterly Journal of Economics* 118, 73-105.

Bacharach, M., Guerra, G. and Zizzo, D.J. (2007). The self-fulfilling property of trust: An experimental study, *Theory and Decision* 63, 349-388.

Battigalli, P. and Dufwenberg, M. (2005). Dynamic psychological games. Manuscript, University of Arizona.

Battigalli, P. and Dufwenberg, M. (2007). Guilt in games. *American Economic Review, Papers and Proceedings* 97, 170-176.

Becker, G. (1974). A theory of social interactions. *Journal of Political Economy* 82, 1063-1093.

Berg, J.E., Dickhaut, J. and McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior* 10, 122-142.

Bénabou, R. and Tirole, J. (2006). Incentives and prosocial behavior." *American Economic Review*, 96(5): 1652-1678.

Bernheim, B.D. (1994). A theory of conformity. *Journal of Political Economy* 102, 841-877.

Charness, G. (2004). Attribution and reciprocity in an experimental labor market. *Journal of Labor Economics* 22, 665-688.

Charness, G. and Dufwenberg, M. (2006). Promises and partnership. *Econometrica* 74, 1579-1601.

Camerer, C. (2003). *Behavioral Game Theory*. Princeton: Princeton University Press.

Cox, J.C. (2004). How to identify trust and reciprocity. *Games and Economic Behavior,* 46, 260-281.

Cox, J.C. and Deck, C.A. (2005). On the nature of reciprocal motives. *Economic Inquiry*, 43, 623-635.

Cox, J.C., Friedman, D., and Sadiraj, V. (forthcoming). Revealed altruism. *Econometrica.*

Dana, J., Cain, D.M., & Dawes, R.M. (2006). What you don't know won't hurt me: Costly but quiet exit in dictator games. *Organizational Behavior and Human Decision Processes*, 100, 193-201.

Dufwenberg, M. and Gneezy, U. (2000): Measuring beliefs in an experimental lost wallet game. *Games and Economic Behavior* 30, 163-182.

Dufwenberg, M., Gächter, S., and Hennig-Schmidt, H. (2006). The framing of games and the psychology of strategic choice. CeDEx Discussion Paper No 2006-20.

Ellingsen, T., Johannesson, M. (forthcoming). Pride and prejudice: The human side of incentive theory. *American Economic Review*.

Falk, A., Fehr, E., and Fischbacher, U. (2007). Testing theories of fairness – Intentions matter. *Games and Economic Behavior* doi:10.1016/j.geb.2007.06.001.

Falk, A. and Kosfeld, M. (2006). The hidden costs of control. *American Economic Review* 96, 1611-1630.

Fehr, E. and Schmidt, K.M. (2006). The economics of fairness, reciprocity and altruism – Experimental evidence and new theories. S.-C. Kolm and J.M. Ythier (eds.): *Handbook of the Economics of Giving, Altruism and Reciprocity,* vol I, Chapter 8. Amsterdam: Elsevier.

Geanakoplos, J., Pearce, D., and Stacchetti, E. (1989). Psychological games and sequential rationality. *Games and Economic Behavior* 1, 60-79.

Glazer, A. and Konrad, K.A. (1996). A signaling explanation of charity." *American Economic Review* 86, 1019-1028.

Guerra, G. and Zizzo, D.J. (2004). Trust responsiveness and beliefs. *Journal of Economic Behavior and Organization* 55, 25-30.

Harbaugh, W.T. (1998). What do donations buy? A model of philanthropy based on prestige and warm glow. *Journal of Public Economics* 67, 269-284.

Hoffman, E., McCabe, K., Shachat, K. and Smith, V.L. (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior* 7, 346-380.

Hoffman, E., McCabe, K. and Smith, V.L. (1996). Social distance and other regarding behavior in dictator games. *American Economic Review* 86, 653-660.

Ireland, N. (1994). On limiting the market for status signals. *Journal of Public Economics* 53, 91-110.

Rabin, M. (1993). Incorporating fairness into game theory and economics." *American Economic Review* 83, 1281-1302.

Ross, L., Greene, D. and House, P. (1977). The false consensus effect: an egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology* 13, 279-301.

Sobel, J. (2005). Interdependent preferences and reciprocity. *Journal of Economic Literature* 43, 392-436.

Tadelis, S. (2007). The power of shame and the rationality of trust. Manuscript. Haas School of Business, UC Berkeley.

Soetevent, A.R. 2005. Anonymity in giving in a natural context? A field experiment in 30 churches. *Journal of Public Economics* 89, 2301-2323.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science* 185, 1124-1131.

APPENDIX I: EXPERIMENTAL INSTRUCTIONS IN EXPERIMENT I

The original instructions for Experiment I were in Swedish. This appendix contains a translation of the instructions.

INSTRUCTIONS (elicitation of beliefs among recipients)

Before we proceed with the actual experiment, we want you to guess the outcome of it. In the experiment you will be anonymously paired with another person in another room. The person in room A will decide how to split SEK 120 between him/herself and the person he/she has been paired with in room B. You will be in room B.

Every individual decision in room A will be anonymous towards both other participants and the experimenters (every person in room A will receive an envelope with six SEK 20 bills and go behind a screen and decide how many SEK 20 bills to leave in the envelope; every receiver in room B will then randomly receive one of these envelopes).

We want you to guess how much, on average, of the SEK 120 a person in room A will give to the person in room B. Write your guess and your student identification number on the form marked "guess", which was handed out together with these instructions (your guess must be between SEK 0 and SEK 120). The person among you whose guess is closest to the actual result in the experiment will win SEK 100; the winner will be notified by e-mail next week and can then collect the prize (if more than one person is equally close they will share the SEK 100).

INSTRUCTIONS (dictator game; the same instructions in room A and B)

Thank you for participating in this experiment. In the experiment each of you will be paired with another person in another room. You will not be told who this other person is, neither during nor after the experiment. Excepting one person in room A who will be chosen to be a monitor, there is an equal number of persons in each room (A and B). **This is room A (B).** Every person in room A and Room B has received these instructions. In the experiment every person in room A (except the monitor) will decide how to divide SEK 120 between him/herself and the person in room B with whom he/she has been paired.

The monitor will receive SEK 120. The monitor's task is to take care of the envelopes we will describe soon. Furthermore the monitor shall control and certify that the instructions we now go through were followed.

The experiment runs as follows. Unmarked envelopes corresponding to the number of participants have been put into a box in room A. All of these contain a smaller envelope and a numbered form marked "guess" (these guesses are described below). All the smaller envelopes contain six SEK 20 bills. The monitor hands out an envelope to each person in room A. When the persons in room A have gotten their envelopes they open the large envelope and take out the form marked "guess", but let the smaller envelope remain inside the large envelope. Each person silently reads the guess and thereafter puts the form back into the envelope.

The monitor then asks one person at a time in room A to come forward. The person takes the envelope and goes behind the screen in room A where no one else can see what happens.

Behind the screen, every person in room A has to decide how many bills, if any, to leave in the envelope. The person then pockets the remaining bills. The decision is up to each person in room A. No one else, including those conducting the experiment, will know what decision a particular person makes.

When the person behind the screen has made his/her decision he/she seals the smaller envelope and puts it in the larger envelope which is also sealed. The person then puts the envelope in the box marked “Mail” and leaves the room.

When all envelopes have been handed in the monitor takes the box with envelopes to room B. The monitor takes up an envelope from the box and opens the larger envelope to see the number on the form marked “guess”. The person with that numbers follows the monitor to an adjacent room where the monitor opens the smaller envelope and writes down the contents and then gives the contents to the person in question who then leaves. The monitor continues until all envelopes have been opened. The experiment is then over.

On the form marked ”guess” the person in room B has guessed the outcome of this experiment, that is, how much a person in room A on average will give to a person in room B. This guess was made before these instructions were handed out and without knowing that the person in room A would get to see the guess. The person in room B whose guess is closest to the actual outcome of the experiment will win SEK 100 (the winner will be notified by e-mail next week and can then collect the SEK 100 prize).

APPENDIX II: EXPERIMENTAL INSTRUCTIONS IN EXPERIMENT II

The original instructions for Experiment II were in Norwegian. This appendix contains a translation of the instructions.

INSTRUCTIONS (trustors)

In this experiment there are two rooms. This is room A. Every subject in this room has a matching subject in room B. It is random who this person is. No one, not any person in either room, will at any time get to know the identity of their counterpart in the other room. It is the same number of subjects in each room.

After the instructions have been read, you will receive two envelopes. One of the envelopes has an identification number and contains NOK 50 and a green piece of paper.

You must decide whether to send the 50 note to your matching subject in room B, or keep the money for yourself.

If you decide to send the 50 note, it will be multiplied with five, which means that the person you are matched with will receive NOK 250. This person will then decide which of the following amounts he/she will return to you; {0,50,100,150,200,250}. You send money to the person you are matched with in room B by placing the 50 note in the envelope with an identification number. In this case you should place the green piece of paper in the envelope with no identification number.

If you decide to keep the money, the person you are matched with will not receive anything and he/she will consequently not be able to return anything to you. In this case you should place the 50 note in the envelope without an identification number and the green piece of paper in the envelope with an identification number.

The envelopes with identification number will be collected in a box as soon as everyone in this room has made their decision. The box will then be handed over to a person in a neutral room. This person multiplies the donation with 5 and brings the envelopes to the other room.

The money you have at the end of the experiment is yours.

The subjects in room B will read this instruction: He or she has the same information as you.

For questions, please raise your hands.

Remember that the envelope with an identification number should contain either NOK 50 or a green piece of paper.

INSTRUCTIONS (trustors; elicitation of beliefs)

Every subject in this room has now had the possibility to donate either 0 or NOK 50. The person you are matched with will receive NOK 250 if you donated NOK 50 and 0 if you decided to keep the NOK 50.

The person who receives the donation will then decide which of the following amounts he or she will transfer back to you: {0, 50,100,150,200,250}. You are now asked to answer the following question:

How much do you expect a person who receives NOK 250 will return?

Kindly indicate your choice on the form you receive. The money that subjects in room B actually returns are retained by the experimenter, and everyone in this room that has answered the most common back-transfer receives a price of NOK 50. The prize will be paid at the end of the experiment.

The forms are collected as soon as they are completed by everyone.

INSTRUCTIONS (trustees)

In this experiment there are two rooms. This is room B. Every subject in this room has a matching subject in room B. It is random who this person is. No one, not any person in either room, will at any time get to know who their matching part in the other room is. It is the same number of subjects in each room.

While you read this instruction the subjects in room A face a choice that is described in this instruction (text in italic). This is the same text as the subjects in room A got.

After the instructions have been read, you will receive two envelopes. One of the envelopes has an identification number and contains NOK 50 and a green piece of paper.

You must decide whether to send the 50 note to your matching subject in room B, or keep the money for yourself.

If you decide to send the 50 note it will be multiplied with five, which means that the person you are matched with will receive NOK 250. This person will then decide which of the following amounts he/she will return to you; {0,50,100,150,200,250}. You send money to the person you are matched with in room B by placing the 50 note in the envelope with an identification number. In this case you should place the green piece of paper in the envelope with no identification number.

If you decide to keep the money, the person you are matched with will not receive anything and he/she will consequently not be able to return anything to you. In this case you should place the 50 note in the envelope without an identification number and the green piece of paper in the envelope with an identification number.

The envelopes with identification number will be collected in a box as soon as everyone in this room has made their decision. The box will then be handed over to a person in a neutral room. This person multiplies the donation with 5 and brings the envelopes to the other room.

The money you have at the end of the experiment is yours.

The subjects in room B will read this instruction: He or she has the same information as you.

The subjects in this room will soon receive an envelope. **Do not open this envelope before you are instructed to.** In this envelope there is either NOK 250 or a green piece of paper. If there is a green piece of paper the person in room A that you are matched with decided to keep the NOK 50.

You will also receive a form that was completed by your counterpart in room A. On this form it is indicated what this person expects that a person who receives NOK 250 will return to the sender. After you have read the instruction you will be asked to open the envelope. If it contains NOK 250 you have to decide how much you want to return to the person you are matched with in room A. You can return any of these amounts: {0,50,100,150,200,250}.

Place the money you decide to return in the envelope.

The money you do not place in the envelope, i.e., the money you decide to keep, should be hidden for the other subjects until the envelopes are collected. Your choice must not be revealed to anyone in this room.

If there is no money in the envelope, there is nothing for you to return.

The envelopes will be collected two minutes after you are told to open them.

When the envelopes are collected the experiment is over. Thanks for your participation.

APPENDIX III: EXPERIMENTAL INSTRUCTIONS IN EXPERIMENT III

The original instructions for Experiment III was in Norwegian. This appendix contains a translation of the instructions.

INSTRUCTIONS (player A (principals))

We want to emphasize that real money is at stake. The amount of money you earn in this experiment is paid out anonymously at the end of the session. No one, not the other subjects or the experimenters, will know your choices and how much money you earned in this experiment.

In this experiment there are two rooms. **This is room A**. All subjects in this room are paired with a subject in room B. You are player A and the person you are matched with player B. It is random who this person is. No one in room A or B will know who they are matched with.

You and your partner have two choices. You can decide IN or OUT. Your partner chooses either ROLL THE DIE or NOT ROLL THE DIE.

If you choose OUT, both you and your partner receives NOK 100 each. In that case the choice of your partner in room B has no payoff relevance.

If you choose IN your payoff depends on whether or not your partner chose to ROLL THE DIE or NOT ROLL THE DIE, and the face of the die if he/she chooses to ROLL THE DIE:

- If your partner in room B chooses NOT ROLL THE DIE he/she gets a payoff of NOK 280, while you get nothing.
- If your partner in room B chooses ROLL THE DIE, the payoff you receives depend on the face of the die. If the face shows 1 you receive nothing while your partner receives NOK 200. If the face of the die shows 2,3,4,5 or 6 you receive NOK 240, while your partner receives NOK 200.

The payoffs to A and B are given in this table

	A gets	B gets
A chooses OUT	100	100
A chooses IN, B chooses NOT ROLL THE DIE	0	280
A chooses IN, B chooses ROLL THE DIE and the die shows 1	0	200
A chooses IN, B chooses ROLL THE DIE and the die shows 2,3,4,5 or 6	240	200

The subjects in room B have received the same information as the subjects in room A.

Your partner in room B make their decision before they know whether A (you) choose IN or OUT. But since their choice only has an impact if A chooses IN, we ask them to make their

choice as if their partner in A has decided to choose IN. The die will be rolled after B has decided either to ROLL THE DIE or NOT ROLL THE DIE.

Together with this instruction you have received a form with your identification number. Make your choice on this form. When everyone in this room has made their choice the forms will be collected by the person that distributed the instruction.

Every participant receives NOK 50 as a show up payment in addition to what you may earn in the experiment.

Please raise your hand if you have questions.

INSTRUCTIONS (player A (principals); elicitation of beliefs)

The participants in room B can choose between ROLL THE DIE and NOT ROLL THE DIE. We now ask you to guess the percentage of participants in room B that chooses to ROLL THE DIE.

Be kind and indicate your guess on the form you soon will receive. Those that are within a margin of 5 percentage units of the percentage of B subjects that actually choose ROLL THE DIE will receive a prize of NOK 50. The prize will be paid at the end of the session.

The form is collected as soon as everyone in the room has completed it.

INSTRUCTIONS (player B (agent))

We want to emphasize that real money is at stake. The amount of money you earn in this experiment is paid out anonymously at the end of the session. No one, not the other subjects or the experimenters, will know your choices and how much money you earned in this experiment.

In this experiment there are two rooms. **This is room B**. All subjects in this room are paired with a subject in room A. You are player B and the person you are matched with is player A. It is random who this person is. No one in room A or B will know who they are matched with.

You and your partner have two choices. Your partner can choose IN or OUT. You choose either to ROLL THE DIE or NOT ROLL THE DIE.

If your partner choose OUT, both you and your partner receives NOK 100 each. In that case your choice has no impact on payoffs.

If your partner chooses IN your payoff depends on whether or not you chose to ROLL THE DIE or NOT ROLL THE DIE, and the face of the die if you choose ROLL THE DIE:

- If you choose NOT ROLL THE DIE you get NOK 280, while your partner gets nothing.
- If you choose ROLL THE DIE, the payoffs you receive depend on the face of the die. If the face shows 1 you get NOK 200 while your partner in room A gets nothing. If

the face of the die shows 2,3,4,5 or 6 you receive NOK 200, while your partner in room A receives NOK 240.

The payoffs to A and B are given in this table

	A gets	B gets
A chooses OUT	100	100
A chooses IN, B chooses NOT ROLL THE DIE	0	280
A chooses IN, B chooses ROLL THE DIE and the die shows 1	0	200
A chooses IN, B chooses ROLL THE DIE and the die shows 2,3,4,5 or 6	240	200

The subjects in room A receive the same information as the subjects in room B.

Note that your choice of ROLL THE DIE or NOT ROLL THE DIE has no impact on the payoffs if your partner in A decided OUT. When you make a decision you are asked to choose as if your partner in room A chose IN.

You will soon receive two forms. One of the forms contains your identification number; indicate your choice on this form. The other form is completed by your partner. After your partner had decided either to choose IN or OUT he/she was asked to guess the percentage of subjects in room B that would choose ROLL THE DIE. When A wrote down their beliefs they did not know that B would receive this information. When everyone in this room has made their choices the form will be collected by the person that distributed the instructions.

After the forms have been collected, each of you will be visited by the experimenter in this room. He will give you a die and ask you to roll it. Everyone must roll the die, independent of the choice you made earlier. The reason you all have to roll the die is that this procedure guarantees that no one in room B can infer whether subjects decided to ROLL THE DIE or NOT ROLL THE DIE. If your choice was to NOT ROLL THE DIE the face of the die have no impact on the payoffs.

Every participant receives NOK 50 as a show up payment in addition to what you may earn in the experiment.

Please raise your hand if you have questions.

FIGURES

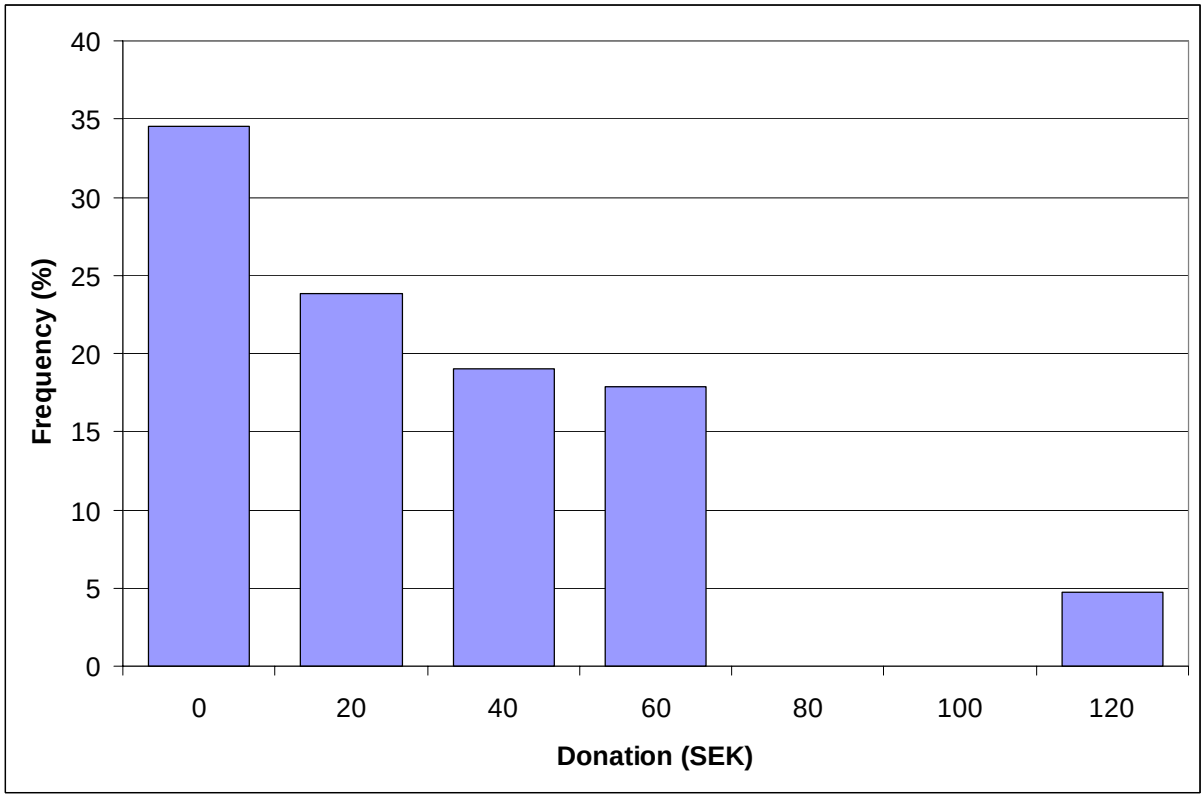

Figure 1. The distribution of donations in Experiment I (the dictator game experiment).

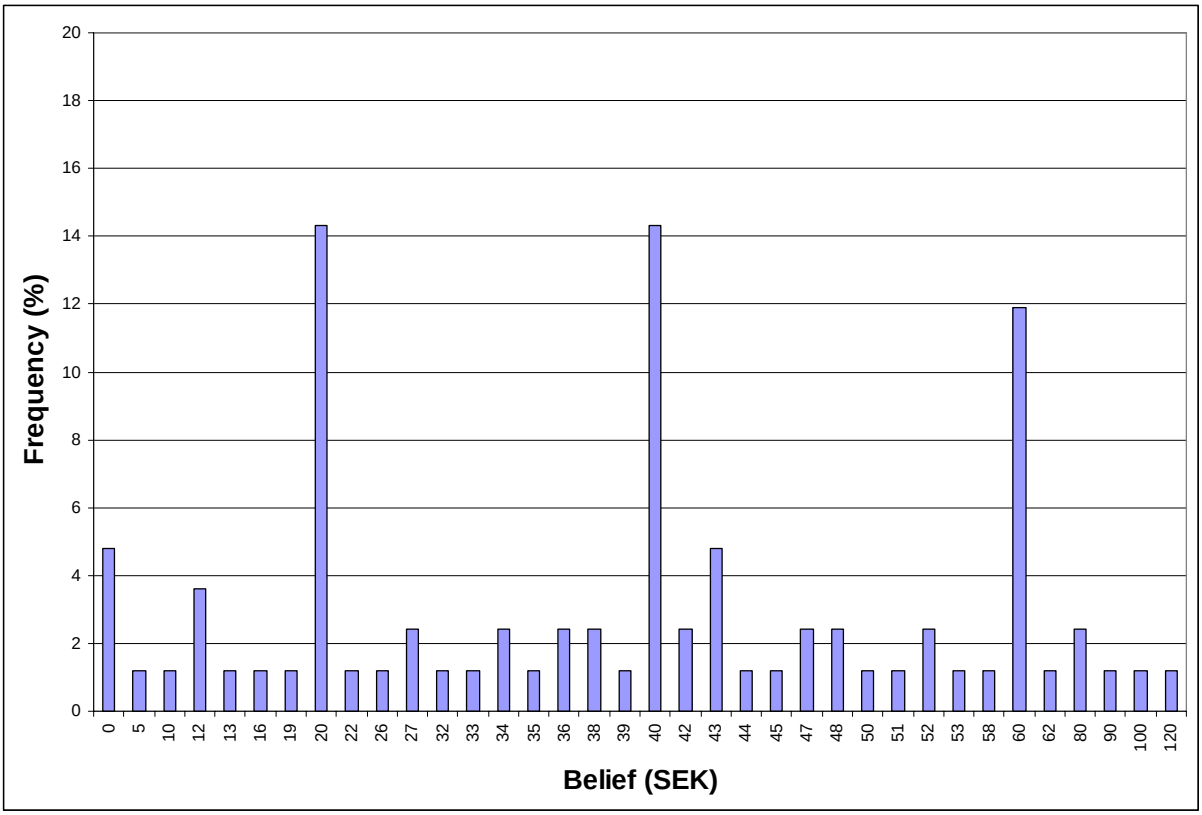

Figure 2. The distribution of beliefs in Experiment I (the dictator game experiment).

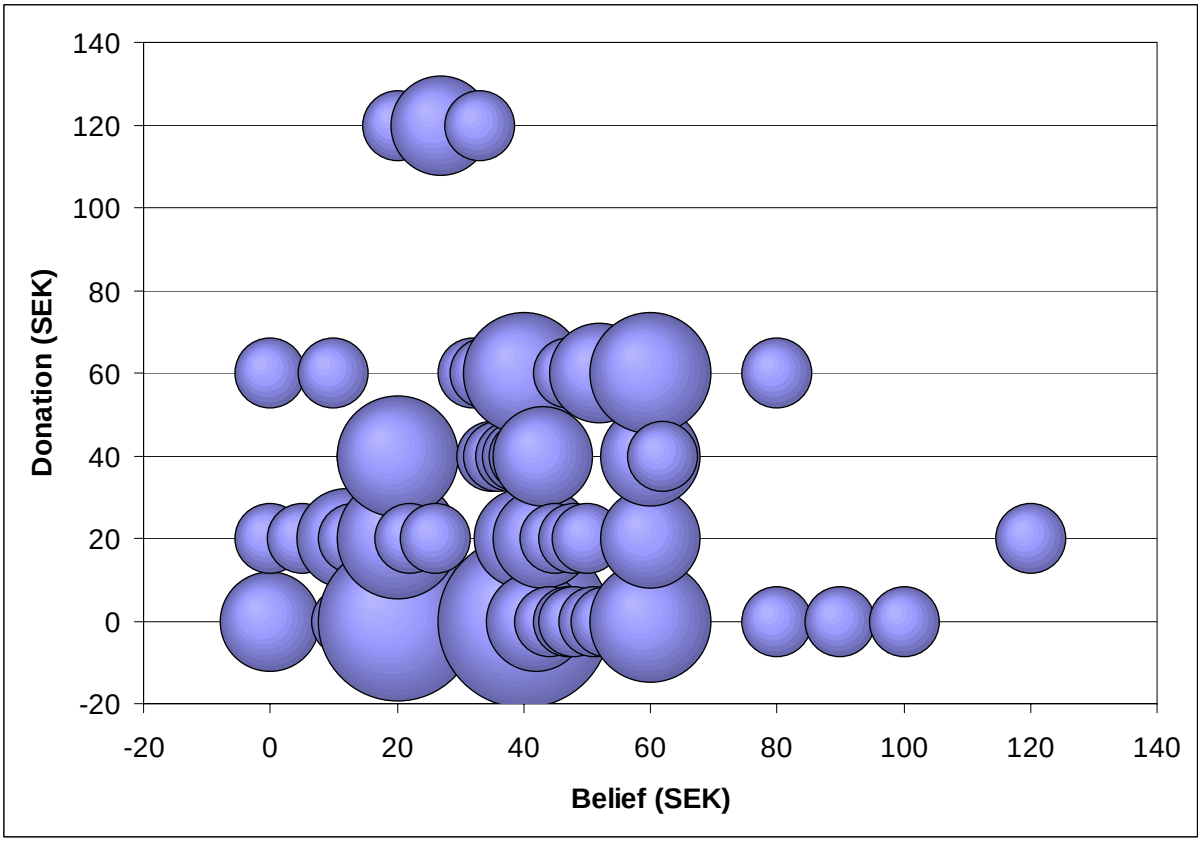

Figure 3. A bubble plot of donations and beliefs in Experiment I (the dictator game experiment).

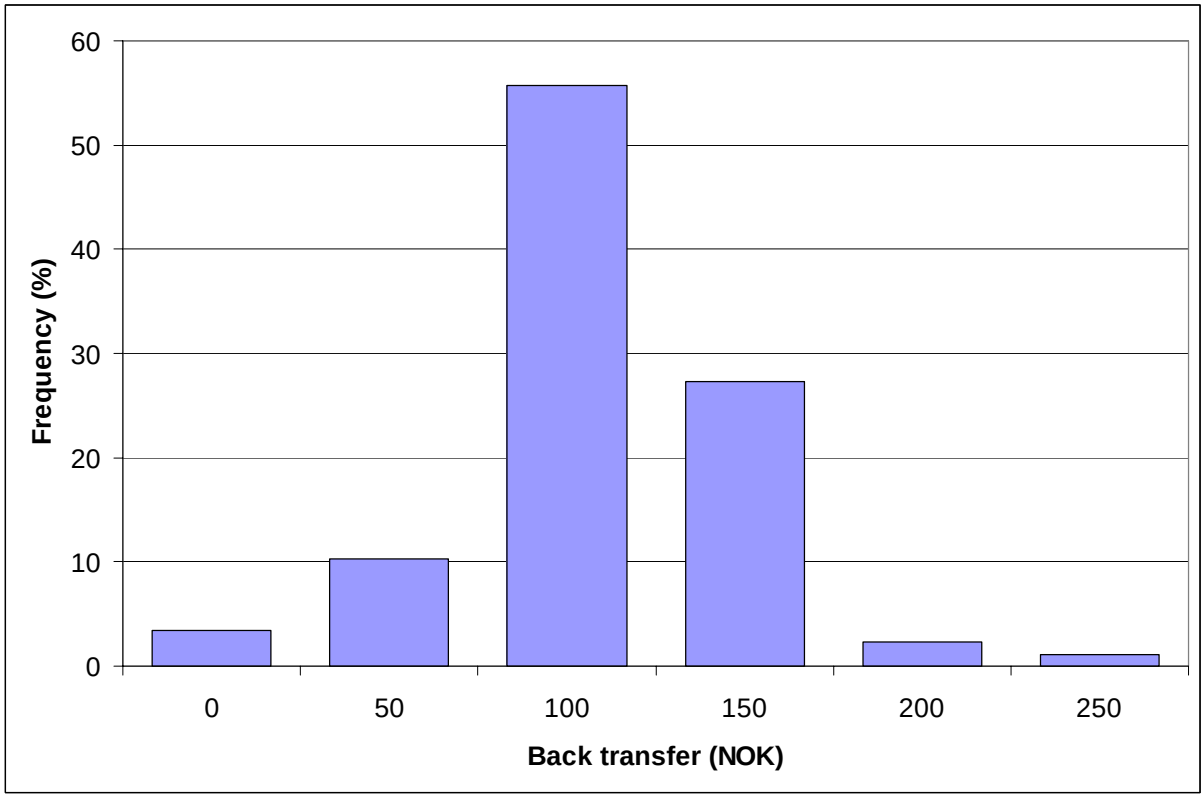

Figure 4. The distribution of trustees' back-transfers in Experiment II (the trust game experiment).

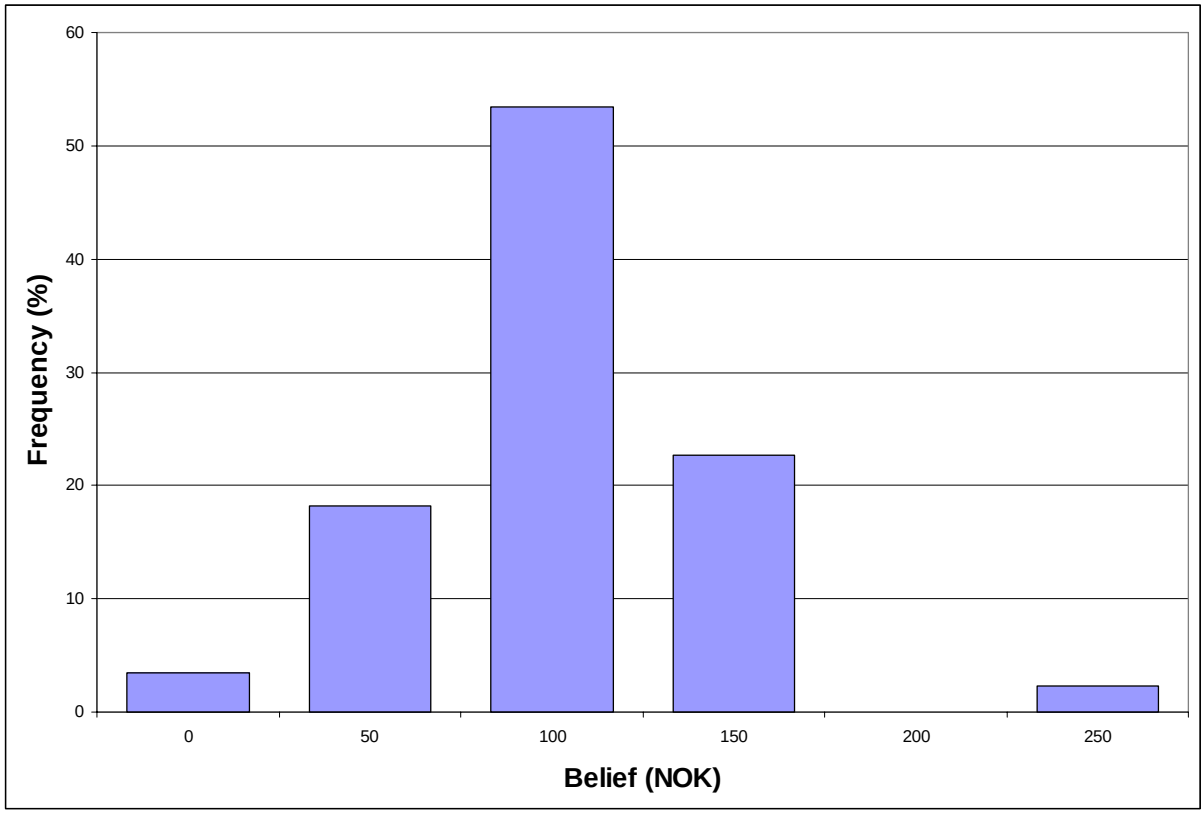

Figure 5. The distribution of beliefs in Experiment II (the trust game experiment).

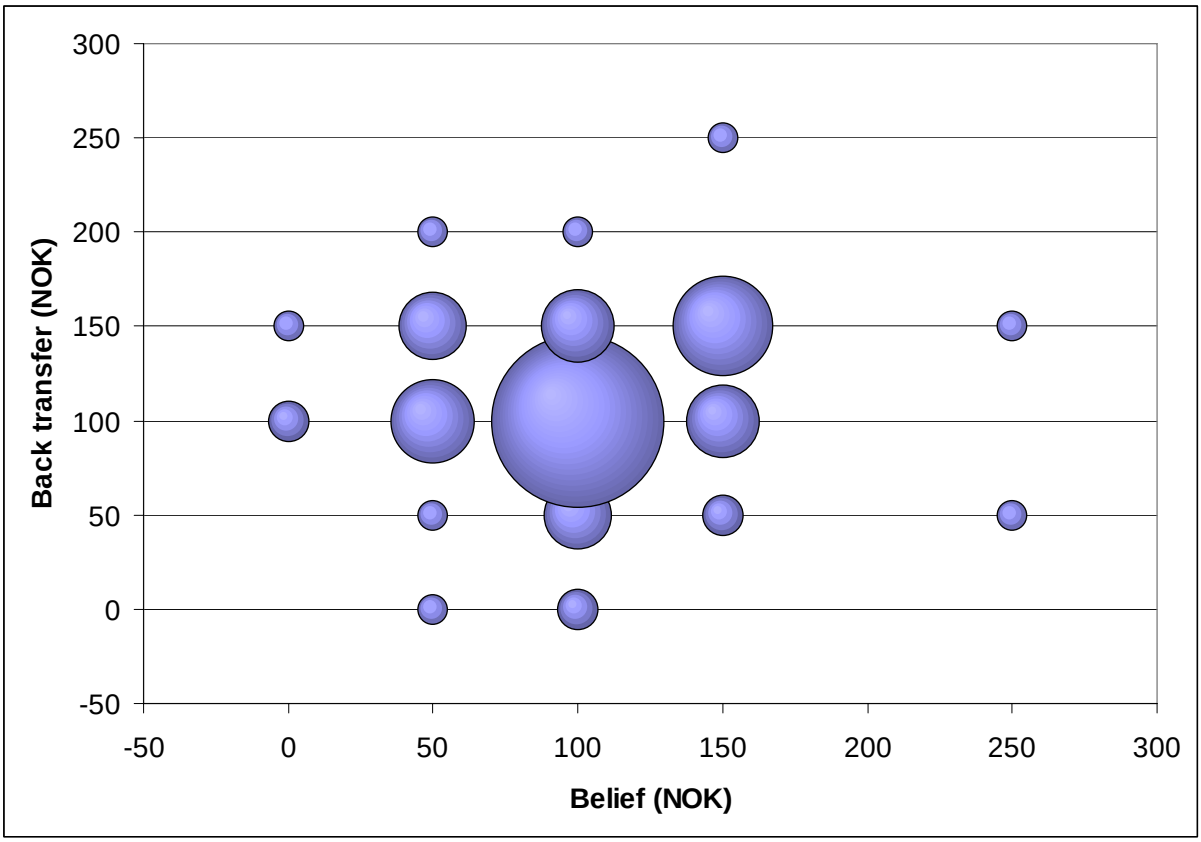

Figure 6. A bubble plot of back-transfers and beliefs in Experiment II (the trust game experiment).

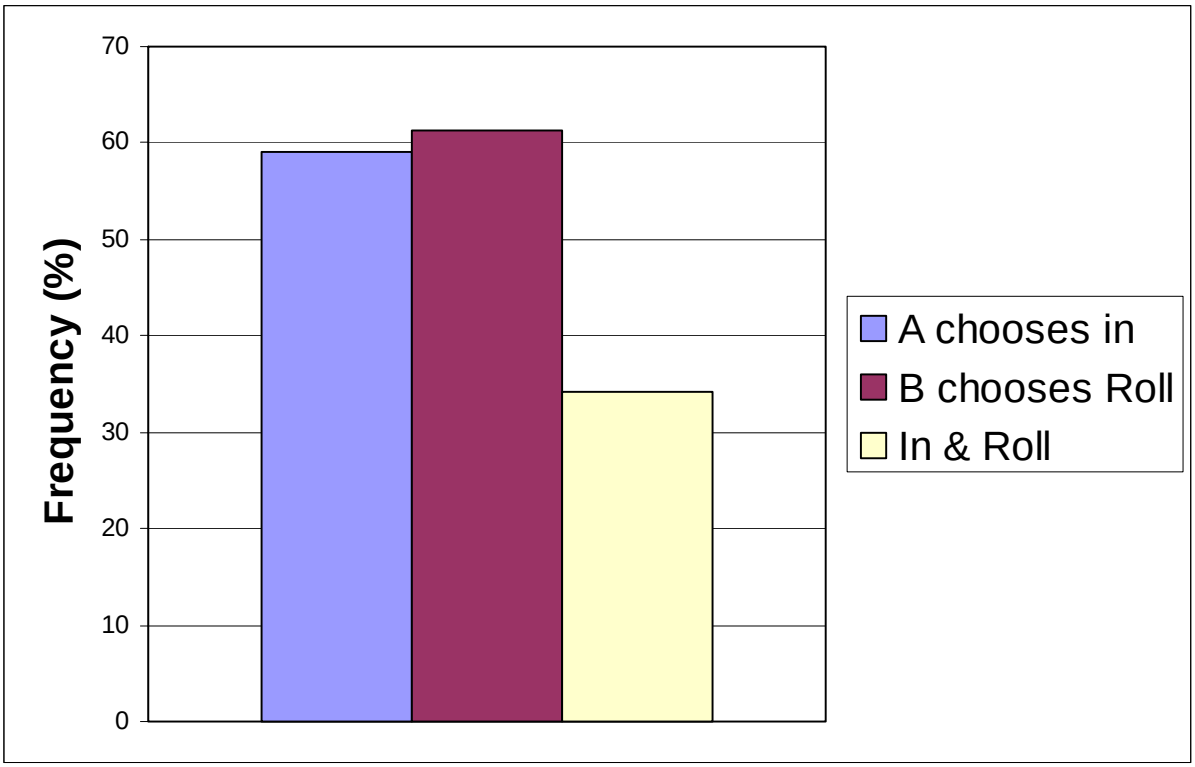

Figure 7. Decisions in Experiment III (the trust game with hidden action experiment).

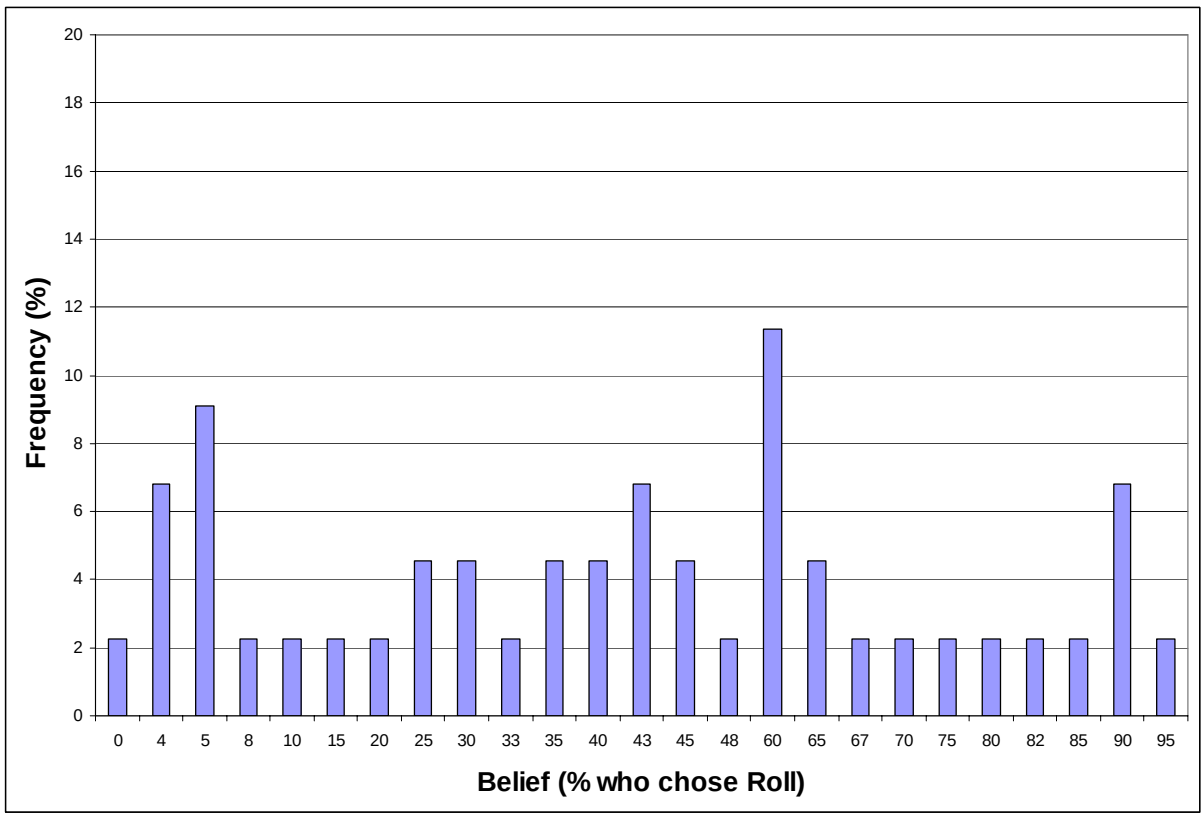

Figure 8. The distribution of beliefs in Experiment III (the trust game with hidden action experiment).

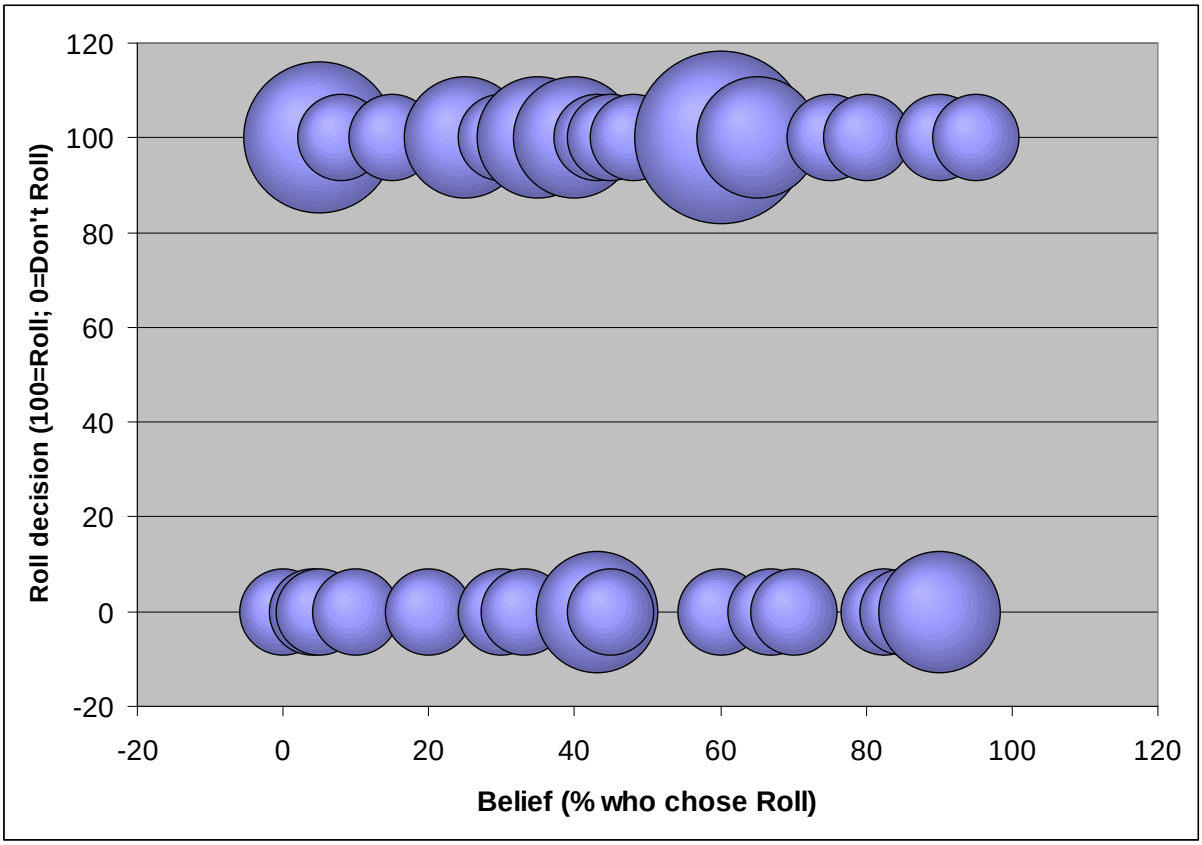

Figure 9. A bubble plot of player B's Roll decisions and beliefs in Experiment III (the trust game with hidden action experiment).