Bare Promises: An Experiment

Gary Charness & Martin Dufwenberg

January 20, 2010

Forthcoming, *Economics Letters*

Abstract: Many individuals dislike making statements that are not true, so non-binding promises may gain power to foster trust & cooperation. If truth-value of a statement is what lying aversion is all about, then these effects should kick in even with bare and impersonal promises. Theoretical accounts of cost-of-lying (implicitly) presume this. We test the idea in a trust game, and report data that provide only limited support for a truth-value based cost-of-lying. We also test for guilt aversion and again find only limited support.

Keywords: Communication, promises, cost-of-lying, truth-value, credibility, guilt aversion

JEL codes: A13, C91, D03

Acknowledgements: We thank Andreas Blume, Vince Crawford, Tore Ellingsen, Håkan Holm, Navin Kartik, David K. Levine, Vai-Lam Mui, Louis Putterman, Matthias Sutter, Christoph Vanberg, a referee, and participants in a seminar at the University of Central Florida and at the 2007 Arne Ryde Symposium on Communication in Games and Experiments at Lund University for very useful comments. We thank the National Science Foundation for support.

Contact: Gary Charness, Department of Economics, University of California at Santa Barbara, charness@econ.ucsb.edu; Martin Dufwenberg, Department of Economics, University of Arizona and Department of Economics, University of Gothenburg, martind@eller.arizona.edu.

1. INTRODUCTION

Previous research has demonstrated that people have some degree of aversion to lying.[1] An important research topic is concerned with why this is the case. Charness & Dufwenberg (2006) (henceforth, "CD") suggest that this happens because decision-makers dislike hurting others relative to what others expect to get.[2] In that paper and elsewhere, promises have been found to foster trust & cooperation.[3] With communication in the picture such "guilt-averse" decision-makers' preferences over choices may change with what is said, as words move beliefs. CD's experiments support this. In their trust-games-with-communication, second-movers often make colorful statements-of-intent ('promises') to exhibit trustworthy behavior. This feeds self-fulfilling circles of beliefs about beliefs that trust & cooperation will ensue.

However, based on data from a clever design, Vanberg (2008) calls CD's conclusions to question to some degree. He argues that "the effects of promises cannot be accounted for by changes in payoff expectations. This suggests that people have a preference for promise-keeping per se." Ellingsen & Johannesson (2004) were probably first to model this, via a "personal cost of being inconsistent," and Chen, Kartik & Sobel (2008) and Kartik (2008) develop theory around the more general notion that decision-makers have a (belief-independent) cost-of-lying.

At their roots this concept makes reference to the *truth-value* of statements: decision-makers dislike making statements that are false. In this paper, we focus on a specific element of the overall research agenda: Is this truth-value all we need to capture an important aspect of

[1] See for example Gneezy (2005). Brandts & Charness (2003) demonstrate that people dislike being told lies.

[2] To illustrate, in restaurants guilt-averse guests tip in proportion to how much they expect waitresses to expect to get, such as no tip in Italy, a couple of coins in Germany, and 16.5% in New York City ("double-the-tax"). For a theory of guilt aversion that applies to general games, see Battigalli & Dufwenberg (2007). For an early experiment finding support for guilt aversion in trust games, see Dufwenberg & Gneezy (2000).

[3] See e.g. Kerr & Kaufman-Gilliland (1994), Ellingsen & Johannesson (2004), Charness & Dufwenberg (2006), Sutter (2009), and Vanberg (2008).

human motivation, or does the context in which the statement was made matter? As in CD, we propose a design that augments a trust game with a pre-play communication stage and we elicit beliefs to allow us to test for guilt aversion.[4] However, while CD allowed for a full page of rich free-form written communication, in our new treatment the protocol is developed to be as bare as possible subject to being rich enough to allow a second-mover to issue a promise. If cost-of-lying (or preference for promise-keeping) depends only on truth-value, this should still be sufficient to foster trust & cooperation.

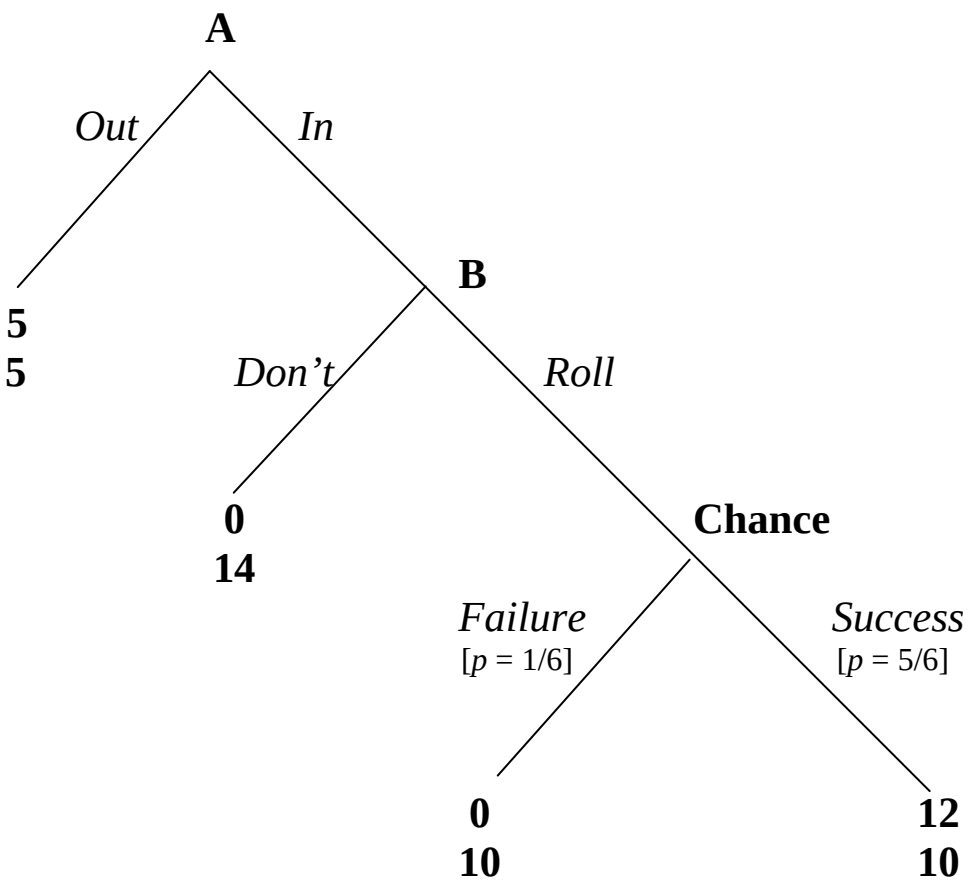
Section 2 describes the design and reports results. Section 3 offers concluding remarks.

2. THE EXPERIMENT

Backdrop

CD's game is reproduced below. The naming of players/strategies reflect experimental protocol. Payoffs reflect dollar payments, not necessarily utilities that may differ with social preferences, e.g. pangs of guilt, or cost-of-lying if there is pre-play communication.

[4] Some controversy has recently arisen regarding the reliability of such elicited-beliefs data. We do not address the issue in this paper; see Ellingsen et al (2009) and Reuben et al (2009) for some evidence and counter-evidence.



CD's benchmark treatment involved no communication and mapped directly to the above game. CD also had a treatment with an opportunity for free-form pre-play communication from B to A; B could print up to a full page of text and send this to A. CD found strong effects on trust & cooperation following a statement-of-intent ("promise"): Comparing post-promise play in the communication-from-B-to-A treatment with play in the no-communication treatment, CD documented significant increases for *In*-rates as well as for *Roll*-rates. Overall, the rate of mutually-cooperative (*In*, *Roll*) strategy profiles was 67% following a free-form promise, far

more than the less than 25% (*In*, *Roll*) rate in the no-communication treatment.[5] These changes in behavior were accompanied by significant changes in beliefs.

In what follows we examine whether a new *bare-promise treatment* generates changes relative to CD's no-communication treatment.

Design

As with CD, sessions were ran at UCSB in a large classroom divided by a center aisle with participants seated at spaced intervals. Our new bare-promise treatment involved three sessions each with 26-36 participants; there were 96 participants in total (no one participating twice). Average earnings were about $14 (including a $5 show-up fee); sessions took about one hour. Participants were referred to as "A" or "B". A coin was tossed to determine which side of the room was A and which was B. Identification numbers were shuffled and passed out face down, and participants were informed that these numbers would be used to determine pairings (one A with one B) and to track decisions for payoffs.[6]

Before playing the game, B could transmit a message to A. Each B was given two sheets of paper. One stated: "I promise to choose *Roll*;" the other was blank.[7] B placed one of the two sheets in an envelope that was conveyed to the appropriate A.

A and B then proceeded to play the game. A first chose *In* or *Out* and then B chose whether to *Roll* or *Don't Roll* a 6-sided die. B made this choice without knowing A's actual choice, but the instructions explained that B's choice would be immaterial if A chose *Out*. As in

[5] CD had many treatments; the rates given here refer to the treatment where the outside option for the first mover was (5,5), which is also the treatment compared in Vanberg (2008).

[6] Complete instructions for the bare-promise treatment are presented in the appendix.

[7] The instructions mentioned that a promise was not binding as otherwise some B's might have felt compelled to choose *Roll*. This approach was also used in Glaeser, Laibson, Scheinkman & Soutter (2000) and Andreoni (2005).

CD, we thus obtain an observation for every B ("the strategy method"). The outcome labeled "Success" in the figure occurred when the die came up 2, 3, 4, 5, or 6 after a *Roll* choice. After the decisions had been collected, a 6-sided die was rolled for each B; this was made clear to the participants in advance, to avoid the anticipated loss of public anonymity for B's who chose *Don't Roll*. This roll was determinative if and only if (*In*, *Roll*) had been chosen.

The outcomes and corresponding payoffs were described to the participants in this chart:

	A receives	**B receives**
A chooses *Out*	**$5**	**$5**
A chooses *In*, B chooses *Don't Roll*	**$0**	**$14**
A chooses *In*, B chooses *Roll*, die = 1	**$0**	**$10**
A chooses *In*, B chooses *Roll*, die = 2,3,4,5, or 6	**$12**	**$10**

A feature of CD's design concerned the provisions made for belief elicitation. We also elicit beliefs in this study. After collecting strategic choices, we passed out guess sheets. A's were asked to guess the proportion of B's who chose *Roll*, conditional on whether or not B sent a promise (see the appendix for instructions). Knowing that A's made this guess, B's were then asked to guess the average guess made by A's who chose *In*, conditional on whether or not B sent a promise. If a guess was within five percentage points of the correct answer, we rewarded the guesser with $2.50 (we also told participants we would pay $2.50 for all B guesses if no A's chose *In*).[8] These guesses constitute the data that we take to represent players' beliefs.

[8] How to best elicit beliefs is a thorny and important issue. We refer to Andersen, Fountain, Harrison & Rutström (2007) for an in-depth discussion about the pros and cons of various methods. Our scheme has the virtue of being simple to describe in instructions (as well as of staying close to CD). As our game is one-shot and we didn't mention guesses until after strategies were chosen, the belief elicitation should not affect participants' prior choices.

Results

A within-bare-promise-treatment comparison of the nature of play with and without promises yields no significant differences for *In*- or *Roll*-rates.[9] However, since only seven B's did not send a promise, the associated statistical tests on our binary-choice data lack power. To check robustness, we therefore compare behavior following a promise in the bare-promise-treatment (where we have 41 observations) with CD's no-communication treatment (where there were 45 observations).[10]

Figure 1 gives a visual impression, and Table 1 provides the full details including the results of difference-in-proportions tests (Glasnapp & Poggio 1985). The lack of support for a trust-enhancing effect of bare promises is clear, as we observe little difference in A's behavior. Regarding the trustworthiness-enhancing effect of bare promises, there is an increase in the *Roll* rate of 16.6 percentage points; this difference is marginally significant on a one-tailed test. Overall, we interpret our results as providing some degree of support for truth-value based cost-of-lying, but the effects are not as large as those seen with CD's free-form promises, where the *Roll* rate after free-form promises was 75%. The observed rate after bare promises is roughly halfway between the rates with no communication and with richer and endogenous promises.

[9] 23/41 A's who received promises chose *In*, compared to 2/7 A's who did not receive promises. 25/41 B's who sent promises chose *Roll*, compared to 3/7 B's who did not send promises.

[10] Of course the baseline for different subject pools may well differ, so that strictly speaking one would prefer a control that is run simultaneously with the treatment. However, all of the sessions were conducted with UCSB students in the same room in all cases. We see little reason to expect a difference in the baseline behavior.

Figure 1 - Bare Promises and Behavior

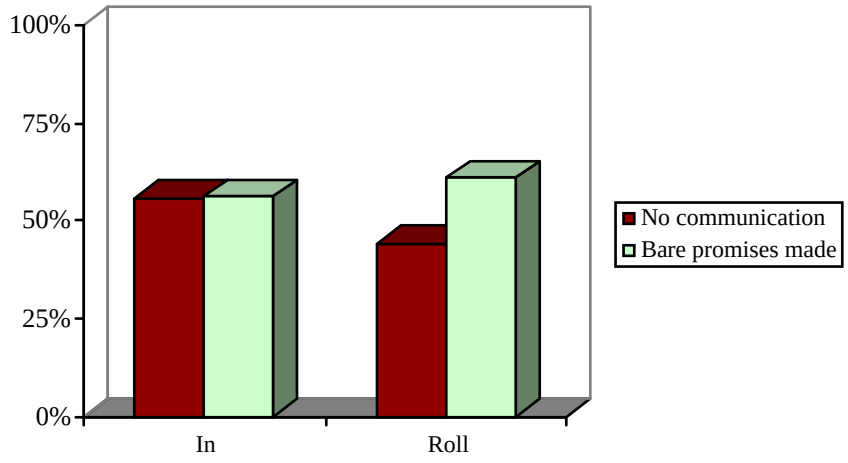


Table 1: Results Across Treatments

	No Communication	Bare Promises Made	Tests of differences
% *In*	25/45 (55.6%)	23/41 (56.1%)	$Z = 0.05$, $p = 0.480$
% *Roll*	20/45 (44.4%)	25/41 (61.0%)	$Z = 1.53$, $p = 0.063$

All tests are one-tailed, presuming an *ex-ante* lying-aversion hypothesis.

We have seen that bare promises have no effect on *In* rates and a marginal effect on *Roll* rates, so that the support for truth-value based cost-of-lying is not overwhelming. Do we find evidence of guilt aversion in our belief data? Recall that the test for guilt aversion is based on B's beliefs concerning A's beliefs. We find that the average belief of B's who choose *Roll* is 60.75, while the average belief of B's who choose *Don't* is 53.05.[11] While the direction of the difference is as predicted, the support for guilt aversion is not as strong here as in CD. The difference in beliefs of B's who choose *Roll* and *Don't Roll* is less than statistically significant (*Z*

[11] We measure the belief of a B-player who sent a promise as his or her guess of the average guess of those A's who chose *In* after receiving a promise. We measure the belief of a B-player who did not send a promise as his or her guess of the average guess of those A's who chose *In* after not receiving a promise

= 0.94, p = 0.174) on a one-tailed Wilcoxon ranksum test, while a simple probit regression of *Roll* against B's beliefs gives Z = 1.12 (p = 0.131, one-tailed) for the coefficient of B's beliefs).

3. CONCLUDING REMARKS

One's regard for the truthfulness of a statement or promise is important for understanding economic interactions involving communication. We test whether violation-of-truth-value can by itself explain why people are averse to lying, by using a design that is rich enough to allow explicit promise-making and is otherwise as bare as possible. In comparison to not permitting communication at all, our data exhibit no effect at all on trust from bare promises, while trustworthiness after bare promises rises to a level intermediate between that observed with no communication and with rich, free-form promises. On balance, and in comparison with the richer promises in CD, these results suggest that bare promises had substantially less effect on behavior.[12] Overall, we find only limited support for truth-value based cost-of-lying; at the same time, we find only limited support for guilt aversion.

While a commitment-based story for promise-keeping may well be relevant, our new results suggest that to explain why people keep promises one needs something more nuanced than concern with truth-value only. It seems clear that context matters; for example, there are situations (such as used car sales, promises made by politicians, tax returns sent to the IRS, and testimony in traffic courts) where lying seems more-or-less expected and liars seem not to suffer

[12] We do not wish to imply that bare communication cannot have large effects. Bare statements of intent have for example been found to be quite effective as an equilibrium-selection device in coordination games (see, e.g., Cooper, DeJong, Forsythe & Ross 1989, 1992; Charness 2000) where players to a large extent have joint interests.

much.[13] Thus, it appears that the issue of why some promises are more effective and credible than others has many shades that will require further study.

REFERENCES

Andersen, S., J. Fountain, G. Harrison & E. Rutström (2007), "Eliciting Beliefs: Theory and Experiments," Working Paper 07-08, Department of Economics, College of Business Administration, University of Central Florida.

Andreoni, J. (2005), "Trust, Reciprocity, and Contract Enforcement: Experiments on Satisfaction Guaranteed," mimeo.

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

Brandts, J. and G. Charness (2003), "Truth or Consequences: An Experiment," *Management Science* **49**, 116-130.

Charness, G. (2000), "Self-serving Cheap Talk and Credibility: A Test of Aumann's Conjecture," *Games & Economic Behavior* **33**, 177-194.

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

Chen, Y., N. Kartik, & J. Sobel (2008), "Selecting Cheap-Talk Equilibria," *Econometrica* **76**, 117-136.

Cooper, R., D. DeJong, R. Forsythe & T. Ross (1989), Communication in the Battle of the Sexes Game: Some Experimental Results," *RAND Journal of Economics* **20**, 568-587.

– (1992), "Communication in Coordination Games," *Quarterly Journal of Economics* **107**, 739-771.

Dufwenberg, M. & U. Gneezy (2000), "Measuring Beliefs in an Experimental Lost Wallet Game", *Games & Economic Behavior* **30**, 163-82.

Ellingsen, T. & M. Johannesson (2004) "Promises, Threats, and Fairness", *Economic Journal* **114**, 397-420.

Ellingsen, T., M. Johannesson, S. Tjøtta & G. Torsvik (2009) "Testing Guilt Aversion", forthcoming in *Games and Economic Behavior*.

Erat. S. & U. Gneezy (2009), "White Lies", mimeo.

Glasnapp, D. & J. Poggio (1985), *Essentials of Statistical Analysis for the Behavioral Sciences*, Merrill: Columbus.

[13] More examples come from a recent study by Erat & Gneezy (2009) on "white lies", which *benefit* another person. They compare situations where the white-liar gains to cases where he or she is hurt, documenting subtle effects.

Glaeser, E., D. Laibson, J. Scheinkman, and C. Soutter (2000), "Measuring Trust," *Quarterly Journal of Economics*, **115**, 811-846.

Gneezy, U. (2005), "Deception: The Role of Consequences," *American Economic Review* **95**, 384-394.

Kartik, N. (2008), "Strategic Communication with Lying Costs," forthcoming in *Review of Economic Studies*.

Kerr, N. & C. Kaufman-Gilliland (1994), "Communication, Commitment, and Cooperation in Social Dilemmas", *Journal of Personality and Social Psychology* **66**, 513–529.

Reuben, E., P. Sapienza & L. Zingales (2009), "Is Mistrust Self-Fulfilling?", *Economics Letters* **104**, 89-91.

Sutter, Matthias (2009), "Deception Through Telling the Truth?! Experimental Evidence from Individuals and Teams", *Economic Journal* **119**, 47-60.

Vanberg, C. (2008), "Why Do People Keep Their Promises? An Experimental Test of Two Explanations," *Econometrica* **76**, 1467-1480.

APPENDIX – INSTRUCTIONS

Thank you for participating in this session. The purpose of this experiment is to study how people make decisions in a particular situation. Feel free to ask us questions as they arise, by raising your hand. Please do not speak to other participants during the experiment.

You will receive $5 for participating in this session. You may also receive additional money, depending on the decisions made (as described below). Upon completion of the session, this additional amount will be paid to you individually and privately.

During the session, you will be paired with another person. However, no participant will ever know the identity of the person with whom he or she is paired.

Decision tasks

In each pair, one person will have the role of A, and the other will have the role of B. The amount of money you earn depends on the decisions made in your pair.

On the designated decision sheet, each person A and person will indicate whether he or she wishes to choose IN or OUT. If A chooses OUT, A and B each receive $5. We will collect these sheets after the choices have been indicated. We will collect these sheets after the choices have been indicated. Next, each person B will indicate whether he or she wishes to choose ROLL or DON'T ROLL (a die). Note that B will not know whether A has chosen IN or OUT; however, since B's decision will only make a difference when A has chosen IN, we ask B's to presume (for the purpose of making this decision) that A has chosen IN.

If A has chosen IN and B chooses DON'T ROLL, then B receives $14 and A receives $0. If A has chosen IN and B chooses ROLL, then B receives $10 and the outcome of a six-sided die determines A's payoff. If the die comes up 1, A receives $0; if the die comes up 2-6, A receives $12. (All of these amounts are in addition to the $5 show-up fee.) This information is summarized in the chart below:

	A receives	**B receives**
Either A or B chooses OUT	**$5**	**$5**
A and B choose IN, B chooses DON'T ROLL	**$0**	**$14**
A and B choose IN, B chooses ROLL, die=1	**$0**	**$10**
A and B choose IN, B chooses ROLL, die=2,3,4,5, or 6	**$12**	**$10**

A Promise

Prior to the decision by A and B concerning IN or OUT, B has an option to promise A that he or she will choose ROLL if A chooses IN. Each B has been given two additional sheets of paper. One sheet has the statement: "I promise to choose ROLL." If you wish to make a promise, please circle this statement. The other is blank, except for the letter B on top. Please return one of these sheets face down to the experimenter, who will convey it to the appropriate A participant, and then A and B will proceed as described above. B may still choose to ROLL or DON'T ROLL after a promise.

A

MAKE A GUESS

We now ask you to guess the percentage of **B's who chose ROLL.**

Please fill in both guesses below.

I guess that _____% of those B's who promised to ROLL actually chose ROLL.

I guess that _____% of those B's who did not promise to ROLL actually chose ROLL.

Payment for the guesses

For each of the two guesses:

If your guess differs by no more than 5 percentage points from the actual percentages, you will receive $2.50.

If your guess differs by more than 5 percentage points from the actual percentages, you will receive $0.

(If there are no B's in one of these two categories, you will be paid $2.50 for your guess in that category, regardless of your answer.)

B

MAKE A GUESS

We have asked A's to make guesses about the percentages of B's who chose ROLL. We now ask you to guess some of the average guesses made **by those A's who chose IN.**

Please fill in both guesses below.

Consider those A's who chose IN after receiving a promise that B would choose ROLL. We asked these A's to guess the percentage of B's who chose ROLL among those B's who promised to ROLL. What do you think is the average guess?

I guess that the average guess is _____%.

Consider those A's who chose IN after not receiving a promise that B would choose ROLL. We asked these A's to guess the percentage of B's who chose ROLL among those B's who did not promise to ROLL. What do you think is the average guess?

I guess that the average guess is _____%.

Payment for guesses:

For each of the two guesses:

If your guess differs by no more than 5 percentage points from the actual percentages, you will receive $2.50.

If your guess differs by more than 5 percentage points from the actual percentages, you will receive $0.

(If there are no A's in one of these two categories, you will be paid $2.50 for your guess in that category, regardless of your answer.)