Promises and Partnership

Author(s): Gary Charness and Martin Dufwenberg

Source: *Econometrica*, Nov., 2006, Vol. 74, No. 6 (Nov., 2006), pp. 1579-1601

Published by: The Econometric Society

Stable URL: https://www.jstor.org/stable/4123084

Econometrica, Vol. 74, No. 6 (November, 2006), 1579–1601

PROMISES AND PARTNERSHIP

BY GARY CHARNESS AND MARTIN DUFWENBERG[1]

We examine experimentally the impact of communication on trust and cooperation. Our design admits observation of promises, lies, and beliefs. The evidence is consistent with people striving to live up to others' expectations so as to avoid guilt, as can be modeled using psychological game theory. When players exhibit such *guilt aversion*, communication may influence motivation and behavior by influencing beliefs about beliefs. Promises may enhance trustworthy behavior, which is what we observe. We argue that guilt aversion may be relevant for understanding strategic interaction in a variety of settings, and that it may shed light on the role of language, discussions, agreements, and social norms in these contexts.

KEYWORDS: Promises, partnership, guilt aversion, psychological game theory, beliefs, trust, lies, social preferences, behavioral economics, hidden action.

1. INTRODUCTION

MUCH OF HUMAN ACHIEVEMENT is produced in partnerships. An extensive body of theoretical research—contract theory—is devoted to understanding which partnerships form, what contracts are signed, and what the economic consequences will be.[2] Considerable attention has been devoted to environments with *hidden action*, where a party's future choice is not contractible. Theorists have shown that if people are rational and selfish (caring only about own income), hidden action is a shoal on which efficient contracting may founder.

We examine experimentally the impact of nonbinding preplay communication on cooperation in a simple one-shot trust game that embodies hidden action. Under conventional assumptions, such communication is ineffective.[3] We explore whether there are psychological aspects that enable communication to promote partnership formation and cooperation.

In particular, building on so-called psychological game theory (see Geanakoplos, Pearce, and Stacchetti (1989), henceforth GPS), we introduce

[1]We are particularly grateful to Ninghua Du for excellent research assistance, and to Paolo Battigalli, Simon Gächter, and several referees for extensive advice. Jon Baron, Jeanette Brosig, Steve Burks, Colin Camerer, Tore Ellingsen, Ernst Fehr, Ayelet Fishbach, Guillaume Fréchette, Dan Friedman, Drew Fudenberg, Uri Gneezy, Brit Grosskopf, David Laibson, Rick Larrick, Dan Levin, David I. Levine, David K. Levine, Tanya Menon, Peter Norman, Matt Parrett, Torsten Persson, David Reiley, Yuval Rottenstreich, David Strömberg, Richard Thaler, Dean Williamson, Bernd Wittenbrink, Bill Zame, and many seminar and conference participants gave us helpful comments. The Swedish Competition Authority and the Russell Sage Foundation provided financial support. Dufwenberg did part of the research while on the faculty at Stockholm University and part while visiting Bonn University and Göteborg University. Thanks to all!

[2]For an entry to the literature, see Bolton and Dewatripont (2005).

[3]To emphasize: we have a *one*-shot game where, under traditional assumptions, the backward-induction solution is unique. We do not consider repeated games in which communication may serve as an equilibrium-selection device.

and test for a new behavioral motivation that furnishes a reason why communication may foster trust and cooperation. The basic idea, which we refer to as *guilt aversion*, presumes that decision makers experience guilt if they believe they let others down. This leads to a nonstandard concept of utility (from the viewpoint of traditional game theory), whereby a player's preferences over strategies depend on his beliefs about the beliefs of others, even if there is no strategic uncertainty. In this connection, messages gain cutting power by shaping beliefs that influence motivation. We examine, in particular, the role of promises in this connection.

The preceding paragraph summarizes in a nutshell the focus of our paper. Our design is primarily conceived neither to test other theories of "social preferences" that may shed light on the role of communication for fostering trust and cooperation nor to compare such theories to guilt aversion. Nevertheless, it is natural to wonder how such theories relate to our data; in Section 5.2 we briefly indicate our take on the subject.

In Section 2, we introduce the game on which our experimental design is based and derive implications of guilt aversion. The experiment is described in Sections 3 and 4. We measure beliefs to enable us to test for guilt aversion. We record messages and examine how "statements of intent" correlate with subsequent choices. With and without communication, the data support guilt aversion. Moreover, it turns out that certain messages, namely promises to perform (statements of intent), inspire a greatly increased level of cooperation.

Section 5 collects a variety of comments about our design and our results. Section 6 sums up and offers concluding remarks regarding the scope for guilt aversion to shed light on the impact of communication in various situations.

2. TRUST, COMMUNICATION, AND GUILT

This section sets the stage for the subsequent experiment. We introduce the trust game on which our design in based (Section 2.1), incorporate communication (Section 2.2), and introduce the key notion of guilt aversion from which our main hypotheses are derived (Section 2.3).

2.1. *A Trust Game with Hidden Action*

We consider trust games like Γ_1 in Figure 1. The names of players and choices anticipate the experimental design. Payoffs are in dollars.

The backward-induction solution for selfish risk-neutral players—strategy profile (*Out*, *Don't Roll*)—is inefficient. Thus, Γ_1 has a "dilemma" flavor like many previously studied trust games.[4] We add a twist: a chance move following (*In*, *Roll*) that determines whether A will get 12 or 0 (with probabilities 5/6 or 1/6). This is essential to the following interpretation.

[4]Compare, e.g., Güth, Ockenfels, and Wendel (1994), Berg, Dickhaut, and McCabe (1995), or Dufwenberg and Gneezy (2000).

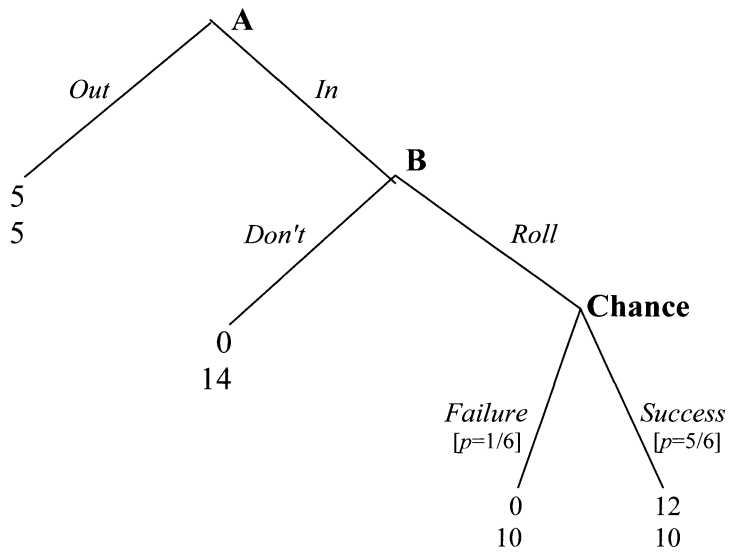

FIGURE 1.—Game Γ_1.

Think of A and B as a principal and an agent: The two consider forming a partnership in which a project is carried out. If no partnership is formed, then no contract is signed, no project is carried out, and the parties each get outside-option payoffs of 5. If the project is carried out, then the contract specifies a "wage" that the principal pays the agent, and a (costly) "effort" that the agent should exert. The project stochastically generates revenue for the principal, the success rate depending on the agent's effort. Strategy profile (*In*, *Roll*) would correspond to a Nash bargaining solution if effort and wage were enforceable. However, the agent's effort is actually unobservable to the principal; the agent is, in practice, free to exert less effort. Moreover, the principal may foresee such a turn of events, dislike it, and refuse to form a partnership. The players' choices *Don't Roll* and *Out* in Γ_1 incorporate these two possibilities.[5]

[5]More precisely, the project can have two outcomes, poor or good. The poor outcome generates revenue 14; the good outcome involves an additional revenue of 12 (so the total revenue is 26). The probability of a good outcome is $5/6 \cdot e$, where $e \in [0, 1]$ is the agent's effort (and $5/6$ may be thought of as her talent). The agent's cost of effort is $4 \cdot e$. Given the outside options of 5 for each party, following Nash (1950) the bargaining solution for risk-neutral and selfish players is the wage-effort pair (w, e) that maximizes $[(14 - w + 5/6 \cdot e \cdot 12) - 5] \cdot [(w - 4 \cdot e) - 5]$. The solution is $(w, e) = (14, 1)$, with resulting payoffs as per (*In*, *Roll*) in Γ_1. If the principal chooses not to join the partnership, each party earns the outside option of 5, as in the end node following *Out*. Otherwise, if the agent chooses $e = 0$, the project fails ($5/6 \cdot 0 = 0$), so the principal gets revenue minus wage equals $14 - 14 = 0$, while the agent gets wage minus effort cost equals $14 - 4 \cdot 0 = 14$, as per (*In*, *Don't Roll*) in Γ_1.

So what about the chance move? Why not replace it with its expected outcome, (10, 10)? The chance move conceptually justifies the given interpretation of unobservability. The principal may get a zero payoff with or without high effort, and he is actually never told which choice the agent made. If, by contrast, outcomes were perfectly correlated with the effort choice, then the agent's choice could arguably be inferred once the payoffs were realized. This would render the unobservability interpretation implausible.[6]

A major issue in contract theory is the choice of contract when a partnership is influenced by hidden action. For example, one may presume that players are selfish and then consider the usefulness of contracts that make the wage contingent on the principal's return. We do not follow this approach. Rather, we stay with a given contract and examine the severity of the problems caused by hidden action in the first place (as implicit in Γ_1; cf. footnote 5).

2.2. *Communication*

We consider treatments that differ according to whether a preplay communication opportunity is present. In the communication treatments, we let one player transmit a message to the other player before they play a game like Γ_1. If the players are selfish, this communication stage obviously has no impact. Words alone cannot change the subsequent payoffs, so (*Out*, *Don't Roll*) remains the unique backward-induction solution.

If other concerns motivate the players, perhaps communication will matter. Several previous studies indeed indicate that communication can affect strategic interaction in one-shot play and thus offer a presumption that communication may matter in our game too. For example, a number of experiments (e.g., Dawes, McTavish, and Shaklee (1977)) provide evidence that face-to-face communication can greatly enhance cooperation in social-dilemma situations.

There are a couple of noteworthy differences between our approach and much of this literature. First of all, previous work deals mainly with prisoner's dilemmas, coordination games, or bargaining games, rather than trust games. Moreover, we do not permit face-to-face communication, but instead use written free-form messages that are transmitted from one party to the other. As Roth (1995) points out, there may be many confounding and uncontrolled effects in face-to-face interaction, and we try to avoid these. Also, by not restricting subjects to a given set of messages, we can study which endogenous messages subjects choose to send.[7]

[6]Independently of the contract-theoretic angle, we note that whether or not B's choice is observable by A may matter to the players' motivation (if they are not selfish). Perhaps B would feel worse choosing *Don't Roll* if he knew that A would know? We do not explore this interesting issue.

[7]Ellingsen and Johannesson's (2004) study is probably the one most closely related to ours. They examine a holdup context that leads to a kind of trust game and they also have written free-

The main novel aspect of our design concerns neither our game nor our communication protocol, however. Rather, it is the particular perspective we provide regarding *why* communication may matter. We discuss this next.

2.3. *Guilt Aversion*

In this paper we focus on guilt aversion, a motivation that provides a route by which communication may influence behavior. Before elucidating the connection with communication, we explain what guilt aversion is and how we test for it.

A guilt-averse player suffers from guilt to the extent he believes he hurts others relative to what they believe they will get. Therefore, he is motivated by his *beliefs about others' beliefs.* Although the idea can be applied to any game, we will focus on Γ_1. Specifically, let $\tau_A \in [0, 1]$ denote the probability that A (initially) assigns to B choosing *Roll*. When B moves, he has a belief (probability measure) regarding τ_A; let $\tau_B \in [0, 1]$ denote its mean. Think of τ_B as a measure of B's belief about A's trust. We use τ_B to define how much B believes he hurts A as well as B's associated guilt:

- If B chooses *Don't Roll*, A gets 0.
- B believes A believes A will get $\tau_B \cdot [(5/6) \cdot 12 + (1/6) \cdot 0] = 10 \cdot \tau_B$.
- The difference, $10 \cdot \tau_B - 0 = 10 \cdot \tau_B$, measures how much B believes he hurts A relative to what A believes she will get, if he chooses *Don't Roll*.
- If B chooses *Don't Roll*, he therefore experiences guilt in proportion to $10 \cdot \tau_B$.

Γ_2 in Figure 2 models this; the parameter $\gamma_B \geq 0$ is a constant that measures B's sensitivity to guilt.

Γ_2 is a nonstandard game in the sense that utilities are not merely numbers at the end nodes. The presence of the belief variable τ_B makes Γ_2, in the terminology of GPS, a *psychological game*. If B is rational, he will choose to *Roll* if $14 - \gamma_B \cdot 10 \cdot \tau_B < 10$. Note that the lower τ_B is, the higher γ_B must be for this inequality to hold, and vice versa.

To derive a testable prediction, we focus on player B. We assume the guilt sensitivity differs among B's and is independent of τ_B.[8] In this case, the higher τ_B is the greater the likelihood that B will choose *Roll*. This is a key research hypothesis in this paper. A test requires us to observe τ_B and our design is set up to achieve this. (We explain how in Section 3.)

---

form messages. For a discussion of other related literature, we refer to their text, which gives a more detailed account.

[8]Tangney (1995) asserts that "there are stable individual differences in the degree to which people are prone to shame and guilt." The hypothesis we derive does not presume that players coordinate on some "equilibrium"; it refers only to the individual player and properties of his/her utility. See Section 5.1 for more comments on this issue.

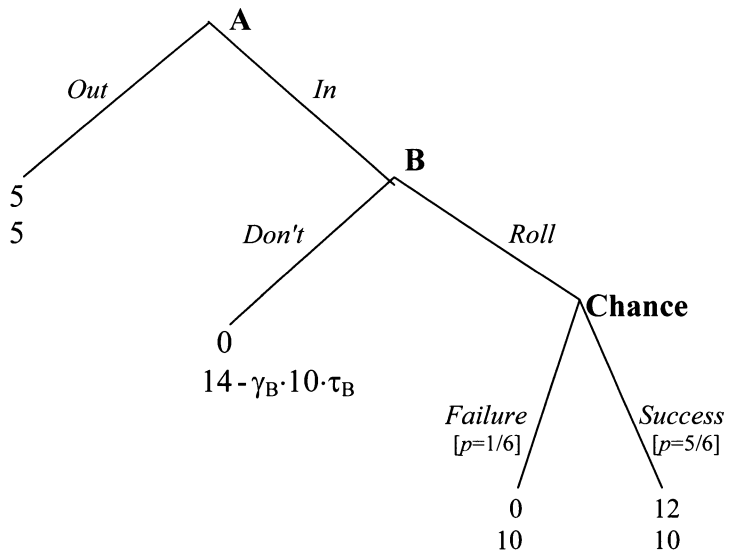

FIGURE 2.—Psychological game Γ_2.

Guilt aversion provides a route by which communication may influence behavior. For example, by making a promise to *Roll*, B may strengthen A's belief that B will *Roll*. This may be plausible, because if B believes A's belief that B will *Roll* is strengthened by the promise, then this strengthens the incentives for B to *Roll* (because the guilt associated with *Don't Roll* goes up). Thus, the promise may lead the parties to play (*In*, *Roll*) rather than (*Out*, *Don't Roll*).[9] Our experiment explores the empirical relevance of this point. In the communication treatments, we observe what messages people transmit, and how this moves beliefs and behavior.

Guilt aversion meshes well with findings in social psychology. See, e.g., Baumeister, Stillwell, and Heatherton (1994, 1995), who (on the basis of autobiographical narratives) suggest that people suffer from guilt if they inflict harm on others. Although guilt could have a variety of sources, one prominent way to inflict harm is to let others down. Baumeister, Stillwell, and Heatherton (1995, p. 173) write that "Feeling guilty [is] associated with... recognizing how a relationship partner's standards and expectations differ from one's own." In economic theory, some applied theoretical work by Huang and Wu (1994) (on

[9]This insight can be compared to some of the ideas explored in the literature on cheap talk in standard games; see Farrell and Rabin (1996) and Crawford (1998) for surveys. However, whereas in a standard game cheap talk may be defined as preplay communication that does not influence the players' evaluation of any given strategy profile, in psychological games such independence cannot be presumed. Talk is cheap only insofar that it does not influence the players' evaluation of strategy profiles for given beliefs.

remorse in corruption) and by Dufwenberg (2002) (on guilt in marriage) considers related ideas for specific trust games.[10] Original to us is the link to communication and the idea that guilt aversion is relevant to general games (see Section 6).

3. DESIGN AND HYPOTHESES

Sessions were conducted at University of California, Santa Barbara, in a large classroom divided in two by a center aisle. Participants were seated at spaced intervals. We had 15 sessions—three each of five treatments—with 24–36 participants per session. No one could participate in more than one session. Average earnings were $16 (including a $5 show-up fee); sessions took about one hour.

The experimental instructions are available as a supplement to the present paper (Charness and Dufwenberg (2006)) In each session, participants were referred to as A or B (as in the games of Section 2). 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.

In our first two treatments, we used exactly the game parameters displayed in Γ_1. In our first treatment, no messages were permitted. In the second treatment, each B had the option to send a nonbinding message to A prior to A's choice of *In* or *Out*. All B's were given a sheet of paper, but could decline to send a message by circling the letter B at the top of the otherwise-blank sheet. Then the B's messages were transmitted to the respective A's before the choice of *In* or *Out*.

Next, B chose whether to *Roll* or *Don't Roll* a six-sided die. Participant B made this choice without knowing A's actual choice *In* or *Out*, but the instructions explained that B's choice would be immaterial if A chose *Out*.[11] We thus obtain an observation for every B. The outcome that corresponds to a successful project occurred only if the die came up 2, 3, 4, 5, or 6 after a *Roll* choice. After the decisions had been collected, a six-sided die was rolled for each B; this was made clear to the participants in advance, to avoid the anticipated loss

[10]Guilt aversion has not previously received much attention by experimentalists. Dufwenberg and Gneezy (2000) and Bacharach, Guerra, and Zizzo (2001) collect data on second-order beliefs and test a related hypothesis, but do not examine communication. Hannan, Kagel, and Moser (2002) and Charness and Rabin (2005) consider the impact of requests or expressed hopes on responder behavior, but do not elicit beliefs or consider guilt aversion.

[11]Although somewhat controversial, this *strategy method* (Selten (1967)) is used extensively in experimental economics and may be best suited to games with few decision nodes. We are not aware of any case where a treatment effect found using the strategy method is not found when using the direct-response method.

of public anonymity for the B's who chose *Don't Roll*. This roll was determinative if and only if (*In*, *Roll*) had been chosen.

Our next two treatments were conducted after observing considerable effectiveness for communication. These treatments used exactly the game parameters displayed in Γ_1, *except* that the payoff vector was $(7, 7)$ rather than $(5, 5)$ in case A chose *Out*. These treatments may be seen as tests of robustness; in this case, the gap between A's expected payoff of 10 after (*In*, *Roll*) and A's reservation payoff is considerably smaller than before, making *In* presumably less attractive to A. Even though communication may be effective when large efficiency gains are available from a successful partnership, perhaps it is ineffective in this case.

Our fifth and final treatment was also conducted after observing the results in the first two initial treatments. Here we use the $(5, 5)$ reservation payoffs of our first two treatments, but change who gets to send the message, so that A sends a message to B.

A critical element of our design involves measuring beliefs (τ_A and τ_B), because these are crucial to the guilt aversion. After we collected the strategic choices, we passed out decision sheets that invited participants to make guesses about the choices of their counterparts and offered to reward good guesses. The A's were asked to guess the proportion of B's who chose *Roll*. Analogously, the B's were asked to guess the average guess made by A's who chose *In*. If a guess was within five percentage points of the realization, we rewarded the guesser with \$5 (we also told participants that we would pay \$5 for all B guesses if no A's had chosen *In*). These guesses represent our measurement of τ_A and τ_B.[12]

We tested several research hypotheses: First, in relation to the relevance of guilt aversion, are *Roll* choices more common when (our measure of) τ_B is high? Second, concerning the role of communication, are *In* and *Roll* choices more common in the message treatments, and is this coincident with higher τ_A and τ_B values, as the guilt-aversion hypothesis would suggest? Third, about the content of the message, do promises or statements of intent play a special role in moving the frequency of *In* and *Roll* choices, and τ_A and τ_B values?

[12]The incentives provided are not the same as those under the alternative of the (more complicated) quadratic-scoring rules (for example, our method excludes guesses less than 5% or greater than 95% as rational responses). We did not ask A's to guess the probability the paired B would choose *Roll*, because we do not observe this likelihood. The observed binary choice would make this simply a *yes* or *no* guess. As regards B's guess, the chosen format is somewhat problematic in the communication treatment in that guilt aversion entails a statement about B's beliefs about the A with whom he is paired, not the average guess of all A's he may be paired with, and a message may influence this belief. We implicitly assume that B's feel that all A's sent similar messages. Overall, we chose our belief-elicitation protocol mainly because it is simple and rather easy to describe in instructions, and also avoids the binary-choice problem. Our idea is to get a rough-but-meaningful ballpark estimate of participants' degrees of beliefs. As our game is one-shot and we did not mention guesses until after strategies were chosen, the belief elicitation should not affect participants' prior choices.

We find it natural to focus on such messages, because previous work has indicated that promises can induce commitments to cooperate (cf., e.g., Kerr and Kaufman-Gilliland (1994), Ellingsen and Johannesson (2004)).

4. RESULTS

We consider the effect of communication in Section 4.1, beliefs and behavior in Section 4.2, and the effect of promises on beliefs and behavior in Section 4.3.

4.1. *The Effect of Communication*

Figure 3 summarizes choices with and without B messages in our payoff calibrations. In the (5, 5) treatment without B messages, 20 of 45 (44%) B's chose Roll and 25 of 45 (56%) A's chose *In*. When B could send a message to A, we observe considerably more cooperation: 28 of 42 (67%) B's chose Roll and 31 of 42 (74%) A's chose *In*. The (*In*, *Roll*) profile occurred 20% of the time (9 of 45 pairs) without communication, compared to 50% (21 of 42 pairs) with messages possible from B's.

We observe similar effects in the (7, 7) treatment. Without B messages, 12 of 48 (25%) B's chose Roll and 11 of 48 (23%) A's chose *In*. When B could send a message to A, we once again observe considerably more cooperation: 24 of 49 (49%) B's chose Roll and 23 of 49 (47%) A's chose *In*. The (*In*, *Roll*) profile

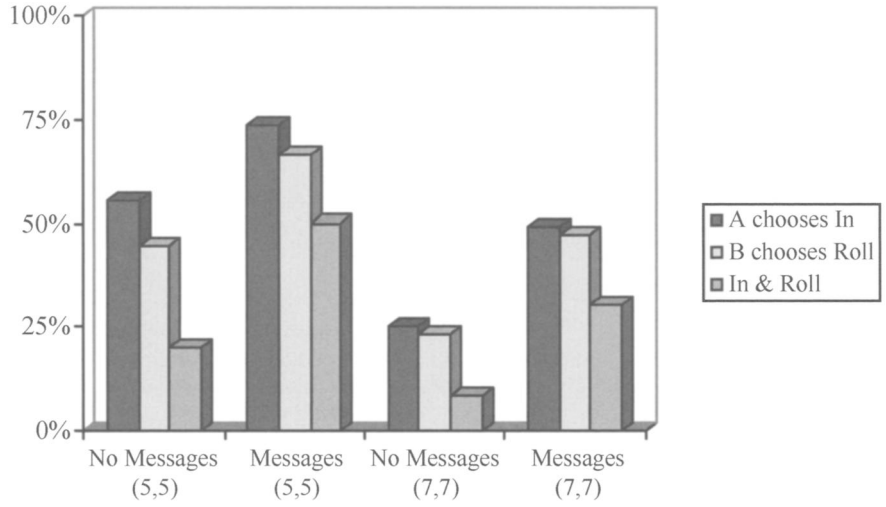

FIGURE 3.—The effect of messages from B.

TABLE I

TESTS FOR THE EFFECT OF COMMUNICATION[a]

	A's *In* Rate			B's *Roll* Rate			(*In*, *Roll*)		
Treatment	M	NM	Z Stat	M	NM	Z Stat	M	NM	Z Stat
(5, 5)	31/42	25/45	1.78**	28/42	20/45	2.08**	21/42	9/45	2.94***
B Messages	(74%)	(56%)		(67%)	(44%)		(50%)	(20%)	
(7, 7)	23/49	11/48	2.48***	24/49	12/48	2.44***	15/49	4/48	2.76***
B Messages	(47%)	(23%)		(49%)	(25%)		(31%)	(8%)	
(5, 5)	31/46	25/45	1.16	18/46	20/45	−0.51	12/46	9/45	0.69
A Messages	(67%)	(56%)		(39%)	(44%)		(26%)	(20%)	

[a]M/NM means that messages/no messages were feasible. The Z stat reflects the test of proportions for the two populations (see Glasnapp and Poggio (1985)). ** and *** indicate $p < 0.05$ and 0.01, respectively, one-tailed tests. Note that the NM data from the (5, 5) case are used as the control in both the first and third rows.

occurred 8% of the time (4 of 48 pairs) without communication, compared to 31% (15 of 49 pairs) with messages possible from the B's.[13]

We can perform formal tests of the null hypothesis that the possibility of communication will not affect behavior by using the aggregate data provided in this subsection; our alternative hypothesis is that communication will improve rates of cooperative behavior.[14] Table I summarizes the effect of communication on behavior for each of our three message treatments.

We can reject the null hypothesis in favor of the alternative for both A's and B's whenever the communication takes the form of messages from B to A. We may conclude that B messages have a major influence on behavior and outcomes in this case.

We mentioned earlier that we also conducted a (5, 5) treatment in which A could send a message to B. In this case, communication was ineffective in improving the rate of cooperative behavior: with communication 31 of 46 (67%) A's chose *In*, while 18 of 46 (39%) B's chose *Roll*; the (*In*, *Roll*) choice occurred 26% of the time (12 of 46 pairs). None of these rates differs substantially or significantly from the rates found in the (5, 5) no-communication treatment, although A's are slightly more likely to choose *In* when A messages are permitted.

[13]Figure 3 also illustrates that "cooperative" choices are more frequent when the available outside option is (5, 5), rather than (7, 7). This is hardly surprising for A, who is taking a bigger chance by choosing *In* when the outside option is (7, 7), but perhaps more unexpected for B. The difference in behavior across payoff calibrations is quite significant for A behavior, both with and without communication ($Z = 2.60$ and $Z = 3.22$, respectively, both significant at $p < 0.010$) and at least marginally significant for B behavior, both with and without communication ($Z = 1.97$ and $Z = 1.70$, respectively, significant at $p = 0.049$ and $p = 0.089$ on two-tailed tests).

[14]Nearly all our tests are conducted using nonparametric statistics. However, we also run probit regressions, which produce essentially the same conclusions. These regressions are available on request.

TABLE II

BELIEFS AND BEHAVIOR[a]

Treatment	A's Average Guess			B's Average Guess		
	In	*Out*	Z Statistic	*Roll*	*Don't*	Z Statistic
(5, 5) no messages	51.3	28.2	2.55***	54.2	39.6	1.99**
(5, 5) B messages	65.4	42.5	2.02**	73.2	45.1	3.20***
(5, 5) A messages	56.7	35.4	2.65***	69.6	50.0	2.80***
(7, 7) no messages	35.7	31.8	1.06	69.4	41.7	3.08***
(7, 7) B messages	70.0	45.3	3.00***	66.9	36.9	3.52***

[a]The Z statistic reflects the Wilcoxon–Mann–Whitney rank sum test for the two populations compared (see Siegel and Castellan (1988)). *, **, and *** indicate $p < 0.10, 0.05$, and 0.01, respectively, one-tailed tests.

4.2. *Beliefs and Behavior*

We have seen that communication affects behavior. Although this is interesting by itself, a key issue motivating our experimental design lies in the relationship between beliefs and choices. Specifically, guilt aversion predicts a positive relationship between B's second-order beliefs (τ_B) and the likelihood that B will choose *Roll*, contrary to the null hypothesis of no relationship between B's beliefs and behavior. If A's respond to incentives, we would also expect that A's who expect B's to be more likely to *Roll* (i.e., A's whose τ_A is higher) will be more likely to choose *In*. Table II details the observed relationship between (measured) beliefs and behavior in each of our treatments.

We observe a strong correlation between beliefs and behavior, both for A's and B's. In each of the five treatments, A's who chose *In* made higher average guesses about the likelihood of *Roll*; in four of these cases, the difference is statistically significant. Results for B behavior are even stronger: In all five treatments, B's who chose *Roll* made significantly higher guesses about A's guesses than did B's who chose *Don't Roll*. Thus, the null hypothesis is strongly rejected, because we find that a B who chooses *Roll* makes a substantially and significantly higher guess about A's guess than a B who chooses *Don't Roll*. We conclude that the support for guilt aversion is considerable in all of our treatments.

Finally, we note that not only do messages from agents increase the probability that B's choose *Roll*, these messages also significantly increase both A's beliefs about this probability and B's beliefs about A's beliefs. The mean guess for A's increased from 41.01 to 59.02 in the (5, 5) treatments and from 32.67 to 56.87 in the (7, 7) treatments; B's mean guess increased from 46.08 to 63.83 in the (5, 5) treatments and from 44.43 to 55.24 in the (7, 7) treatments.[15]

[15]The Wilcoxon–Mann–Whitney test gives $Z = 2.82, 4.07, 2.65$, and 1.77 for the respective comparisons. One-tailed tests yield significance at $p = 0.002, 0.000, 0.004$, and 0.038, respectively.

4.3. *Promises*

We have seen that beliefs differ substantially for people who choose different actions. Is there some particular aspect to messages that causes these beliefs to be so affected by communication? We focus on whether or not a message contains a statement of intent or promise. Because messages can have nearly any form, this requires a classification of the messages. We use three rough categories: promises, empty talk, and no message; our classification is given in the Appendix, along with the raw data on individual choices.[16] Promises are only meaningful when they come from B (although one A nevertheless promised to choose *In*!), so we need only consider the two B-message treatments. The promises category is broad, including any statement of intent that we found. To be sure, some messages were on the boundary between promises and empty talk, and could arguably be placed in either category; nevertheless, the overall pattern is quite clear and is robust to alternative classifications.

The null hypothesis in this case is that statements of intent will not affect behavior, whereas our alternative hypothesis is that such statements will make cooperative behavior more likely. Table III shows A and B behavior according to whether a promise was sent or received.

In all cases but one, the *In* rate, the *Roll* rate, and the ex post (*In*, *Roll*) realizations were much higher following a promise than otherwise. Note that N is fairly small here, because we split the observations in each treatment into

TABLE III

PROMISES AND BEHAVIOR[a]

	A's *In* Rate			B's *Roll* Rate			(*In*, *Roll*)		
Treatment	P	NP	Z Stat	P	NP	Z Stat	P	NP	Z Stat
(5, 5)	22/24	9/18	3.04***	18/24	10/18	1.32*	16/24	5/18	2.49***
B messages	(92%)	(50%)		(75%)	(56%)		(67%)	(27%)	
(7, 7)	16/24	7/25	2.71***	20/24	4/25	4.71***	14/24	1/25	4.13***
B messages	(67%)	(28%)		(83%)	(16%)		(58%)	(4%)	
Pooled	38/48	16/43	4.07***	38/48	14/43	4.49***	30/48	6/43	4.73***
	(79%)	(37%)		(79%)	(33%)		(62%)	(14%)	

[a]P/NP means that a promise/no promise was sent or received. The Z stat reflects the test of proportions for the two populations compared. *, **, and *** indicate $p < 0.10, 0.05$, and 0.01, respectively, one-tailed tests.

[16]It is common in social psychology to code responses according to various classifications. While we only consider the classification in the text, the complete messages are available on the *Econometrica* supplementary materials website for those readers who wish to consider alternative coding (Charness and Dufwenberg (2006)). In the Appendix to this article we provide all messages for one particular treatment ((5, 5) messages from B). Some of the messages are rather colorful and serve well to enliven proceedings in seminars. Consider, e.g., message 7 in session 3 of Table A.I, which contains a poem by Samuel Francis Smith and fictitious references to desires and advice from some famous persons.

TABLE IV

PROMISES AND BELIEFS[a]

	Average A Guess			Average B Guess		
Treatment	P	NP	Z Stat	P	NP	Z Stat
(5, 5)	65.8	50.0	1.63*	66.2	59.9	1.10
B messages	(24)	(18)		(24)	(18)	
(7, 7)	63.1	50.9	1.44*	59.6	51.0	1.17
B messages	(24)	(25)		(24)	(25)	
Pooled	64.4	50.5	2.24**	63.1	54.7	1.74**
	(48)	(43)		(48)	(43)	

[a]P/NP means that a promise/no promise was sent or received. The number of observations is in parentheses. The Z stat reflects the Wilcoxon rank sum test for the two populations. * and ** indicate $p < 0.10$ and 0.05, respectively, one-tailed tests.

two categories; if we compensate for this by pooling the data from the two treatments, the differences in behavior are even more significant. Thus, we can strongly reject the null hypothesis in favor of the alternative.

Regarding statements of intent and beliefs, the null is that there is no relationship, while the alternative hypothesis is that guesses will be higher with promises. Table IV shows average A and B guesses in the B-message treatments, according to whether a promise was sent or received.

In every case, guesses are highest when a promise is made, but no within-treatment test is more than marginally significant. Once again, the number of observations is fairly small for these tests, and if we pool the data from the two treatments to increase the sample size, we do see results that are significant at $p = 0.05$ on the indicated one-tailed test. Thus, the evidence tends to go against the null hypothesis, with promises affecting beliefs, but the effects are modest: In the pooled B-message treatments, A guesses after promises are 27.5% higher than after nonpromises, while B guesses after promises are 15.4% higher than after nonpromises.[17]

5. DISCUSSION

5.1. *Equilibrium and Learning*

Our primary goal is to test for guilt aversion, not whether people have correct beliefs about one another. Therefore, in Section 2.3, when we derived our key research hypothesis (that our measure of τ_B is correlated with the likelihood

[17]It is interesting to compare our results here with a finding of Glaeser, Laibson, Scheinkman, and Soutter (2000, pp. 821, 830). While we compare the effect of endogenously generated promises within a treatment, they make an across-treatment comparison of the effect of an exogenously specified promise opportunity (in another trust game). Their promise condition seems to anchor responses in accordance with the promise; this accords well with a guilt-aversion hypothesis.

of a *Roll* choice) we did not invoke any equilibrium supposition. If we were to do that, we would run the risk of incorrectly rejecting a valid insight about motivation only because people did not coordinate well. Moreover, given that we run a one-shot experiment, with no chance for learning, it seems a bit extreme to assume that people will be able to make accurate predictions about one another.

One might wonder what would happen if people played recurrently, allowing for learning. While leaving this for future research, we wish to point out a pertinent issue: The theory of learning in traditional games centers on assumptions regarding how players over time observe past strategic choices (or paths); see Fudenberg and Levine (1998). In psychological games, finding a best response may in addition require players to learn about the beliefs about beliefs of others. However, beliefs of other players are not easy to observe.

This raises important questions for experimentation as well as for theory. In experiments, is it possible to control beliefs about other player's beliefs? For example, by having the game played repeatedly round robin, and revealing each round the distribution of player B's choices, one might expect player A's beliefs to converge on actual play over time; if B infers this, then the same would be true for B's beliefs about A's beliefs. If, on the other hand, the distribution of play is not revealed, it may be much harder for players to learn. The theoretical issue in this connection concerns formulating appropriate equilibrium concepts for these different cases.[18]

5.2. *Alternative Theories of Motivation*

Our experiment is designed to allow us to test for guilt aversion, not to pit guilt aversion against alternative theories that may or may not explain our data. Nevertheless, it may be useful at this point to make a few comments about how some alternative models of social preferences (developed in response to experimental evidence indicating that decision makers are often not selfish) relate to our data.[19]

Models of distributional preferences (e.g., Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002)) are not suitable for explaining the impact of communication in Γ_1. Participant B knows the distributional

[18]Geanakoplos, Pearce, and Stacchetti (1989) develop several notions of equilibrium for psychological games. An earlier working paper version of this paper drew on their framework to develop a notion of guilt-aversion equilibrium. We dropped this material here because, as explained, our experimental design is not motivated with reference to any equilibrium. Battigalli and Dufwenberg (2005) argue that for many purposes (possibly including learning) the GPS framework needs to be generalized, and they extend the GPS analysis in several directions (including not assuming equilibrium, having incomplete information, and allowing updated beliefs to influence utility).

[19]For descriptions of the experimental evidence and the social-preferences literature it has inspired, see Fehr and Gächter (2000), Fehr and Schmidt (2002), and Sobel (2005). Guilt aversion is not covered.

consequences when she moves, so whether or not there is preceding communication cannot influence her choice.[20]

Models of "intentions-based reciprocity" (e.g., Rabin (1993), Dufwenberg and Kirchsteiger (2004)) build on psychological game theory and involve belief-dependent utility. Nevertheless, they would have difficulty explaining our result that promises reinforce trust and cooperation. To see why, look at Γ_1. Suppose B promises to *Roll*. If A believes this and if B believes that A believes this, then A comes across as *less* kind (because B realizes that A realizes that B gets a lower material payoff if B chooses *Roll* than if B chooses *Don't Roll*). Thus, B would have less reason to *Roll* than if B did not make the promise. Unlike with guilt aversion, promises, if believed, lead to reasons to renege when players are motivated by intentions-based reciprocity.

Levine (1998) and (part of) Gul and Pesendorfer (2005) develop models of "goodness-based reciprocity" in which (rather than having utilities depend directly on beliefs) players care about the perceived goodness of others' types. These models may be well suited for explaining the impact of communication in games, because they assume incomplete information about preference characteristics and the addition of preplay messages may create new signaling opportunities. However, as developed, these models presume equilibrium play, and issues analogous to those we discussed in Section 5.1 apply. We do not address this further here.

We will discuss one final motivational force at some length: a *fixed cost of lying*. Gneezy (2005) presents evidence that indicates that people do not like to lie. Our take is as follows: A guilt-averse person who lies and thereby influences others' beliefs suffers from guilt when he does not live up to these beliefs. This provides a disincentive to lie and a complementary objective to issue promises so as to gain commitment power in contexts where these statements would be believed. However, there is an alternative: Perhaps people experience a fixed and belief-independent cost of lying.[21] That, too, could explain our results about promises and about the ineffectiveness of A messages.

We now argue against this. First, guilt aversion, but not a fixed dislike of lying, predicts a positive relationship between the likelihood of *Roll* choices and τ_B in Γ_1. Second, a fixed dislike of lying can explain selfless choice only in contexts where lying can occur, whereas guilt aversion is a generally applicable idea. Third, guilt aversion, but not a fixed dislike of lying, admits that in certain contexts decision makers do not suffer if they lie (as long as this is expected).[22]

[20]The nondistributional element of Charness and Rabin (2002) does not apply here, because it would apply only to misbehavior by the first mover, whereas (in Γ_1) B's only respond to favorable plays by A's.

[21]See Ellingsen and Johannesson (2004) for a model that incorporates such assumptions.

[22]One example is poker, where players are (clearly!) expected to maximize their own earnings. A per se dislike of lying is not a factor in poker; leading poker texts actively encourage lies, or at least very deceptive use of language and demeanor. See Brunson (2002, pp. 427–428) for an example. Because no one expects truth-telling, there is no guilt associated with deceitful chit-chat.

Fourth, and finally, we can present a regression that supports guilt aversion over a fixed cost of lying: If the latter effect were present in our data, we might expect a difference in *Roll* behavior across the A-message and B-message treatments in the (5, 5) outside-option case, controlling for B's guess. To test this, we perform a probit regression using only the data from these two treatments, with *Roll* as the dependent variable and dummy variables for A messages and for the interaction between A messages and B guesses (standard errors are in parentheses):

$$\begin{aligned} Roll = {} & -1.924 + 0.027 * \text{Guess} + 0.054 * \text{A\_message} \\ & \quad (0.645)\ (0.010) \qquad\qquad (0.991) \\ & - 0.010 * \text{A\_message} * \text{Guess}. \\ & \quad (0.015) \end{aligned}$$

Participant B's guess is important for B's decision whether to *Roll*, but there is no difference across treatments (reflected in the insignificance of the coefficient of both terms with an A-message dummy). This indicates that, holding beliefs constant, B's in the B-message treatment are no more likely than B's in the A-message treatment to *Roll*, suggesting that a fixed dislike of lying is not a major factor in our data.

5.3. *Do Choices Cause Beliefs?*

It has been suggested to us that something akin to a false consensus effect (cf. Ross, Greene, and House (1977)) might produce a positive correlation between B's second-order beliefs and the likelihood of choosing *Roll*. False consensus usually means that a person believes others would act similarly rather than that a person believes others believe he or she would make a certain choice. The related idea here is that B's would think that other B's choose like them and that A's beliefs lean in this direction too. Hence B's beliefs about A's beliefs would resemble B's choice. This would suggest that choices shape beliefs about beliefs, rather than vice versa. Our design does not allow us to rule out that some effect along these lines affects our data.[23]

6. CONCLUSION

We examine the impact of communication in a one-shot principal–agent game designed to capture the essence of hidden action as treated in contract theory. We find that promises (or statements of intent) sent from agents to principals enhance trust, cooperation, and efficiency.

[23]However, we note that a few recent papers test whether beliefs cause behavior or vice versa in various games, and conclude in favor of the former; see Croson and Miller (2004), Fischbacher, Gächter, and Fehr (2001), and Frey and Meier (2004).

The evidence squares well with a notion of guilt aversion, which implies that the more the agent believes his principal expects to be helped, the more likely the agent is to actually help. This ties in nicely with communication; words may affect the agent's beliefs (about what the principal expects) and so may change the degree of guilt he experiences.[24] A promise may feed a self-fulfilling series of beliefs about actions and beliefs, ultimately rendering the promise credible and so potentially attractive to an agent who may benefit from this commitment device.

Not all forms of communication are efficient in moving beliefs, motivation, and behavior. Neither messages by principals nor nonpromising messages by agents has a positive effect. This is not at odds with guilt aversion. Guilt aversion presumes that motivation is belief-dependent, in a particular way. However, guilt aversion in itself does not suggest which forms of communication move beliefs. In this area, we merely record our findings.

The idea of someone feeling guilty from letting others down extends beyond the trust game we have focused on for the most part. We propose that there are a variety of partnerships where guilt aversion and communication may be relevant. Examples include husband and wife, lawyer and client, procurement agency and contracted firm, inventor and producer, talented young golfer and rich sponsor, co-owners of firms, employer and employee, and cartelists. Beyond shedding light on specific partnerships and the impact of direct communication between the parties, we feel that guilt aversion may play a role in some other regards. We close the paper with some speculative remarks about this that are intended to inspire future research.

First, we propose that the ideas that go into the notion of guilt aversion may help explain subtle aspects regarding how people use language. Why do people discuss, argue, and debate so much? Perhaps they are bargaining on what they should all agree is the right thing to do. Perhaps guilt aversion makes people adhere to agreements once they are made. Perhaps guilt aversion can explain respect for democratic decision making from voters who have accepted the legitimacy of the rules of some political process.

Second, do people manipulate the guilt aversion of others in self-serving ways? For example, do authors of research papers attempt to convey, between the lines, the impression that they expect their paper to be accepted in a good journal? That would make sense if their referees were guilt averse; facing a marginal decision, such a referee may be swayed toward acceptance to avoid the guilt she would experience if she rejected the paper and let the authors down.

A final issue concerns the relationship between guilt aversion and social norms. The literature on social norms is vast (see Elster (1989) for a discussion). One central idea is to view a social norm as a moral expectation, which

[24]This observation invites reflection on the idea that the *framing* of a (psychological) game may affect beliefs, and thus affect motivation and behavior. Dufwenberg, Gächter, and Hennig-Schmidt (2005) explore this idea.

people are inclined to live up to.[25] We suggest that in many cases guilt aversion can provide a form of microfoundation for this. Take the case of tipping as an example. Waiters and waitresses in the United States generally expect a 15% tip; this norm may shape everyone's expectations. Yet, guilt aversion may furnish an underlying motivation for why people behave accordingly. There is a norm, it shapes the server's expectation, and the customer lives up to this expectation because he would feel guilty if he did not.[26]

Dept. of Economics, University of California at Santa Barbara, 2127 North Hall, Santa Barbara, CA 93106-9210, U.S.A.; charness@econ.ucsb.edu, http://www.econ.ucsb.edu/~charness/

and

Dept. of Economics & Economic Science Laboratory, University of Arizona, Tucson, AZ 85721-0108, U.S.A.; martind@eller.arizona.edu, http://www.u.arizona.edu/~martind1/.

Manuscript received December, 2003; final revision received May, 2006.

APPENDIX: MESSAGES

The complete messages are available on the *Econometrica* supplementary materials websites, Charness and Dufwenberg (2006). Here we provide all messages for only one particular treatment, namely the '(5, 5) messages from B' treatment, alongside the subjects' choises and our messages classification.

In the appendix table, P = Promise, E = Empty Talk, N = No Message, R = *Roll*, and DR = *Don't Roll*.

TABLE A.I

(5, 5) MESSAGES FROM B

Sess.	ID	Message	Class	A	B
1	1	Please choose In so we can get paid more.	E	Out	DR
1	2	Choose in, I will roll dice, you are 5/6 likely to get 2,3,4,5, or 6 → $12. This way both of us will win something.	P	In	DR
1	3	If you stay in, the chances of the die coming up other than 1 are 5 in 6 – pretty good. Otherwise, we'd both be stuck at $5. (If you opt out)	E	In	DR

Continues

[25]See Bernheim (1994) and Dufwenberg and Lundholm (2000; note especially footnote 5) for models in this spirit.

[26]This is not to say that expectations regarding tipping in restaurants can never be manipulated. Anecdotal evidence illustrates this point: At the *Crab House* restaurant in San Francisco, guests are given a plastic card that reads (in six languages), "Thank you for dining with us. Many guests ask us about tipping. We want you to know that no additional tip or service charge has been added to your bill. In the United States, quality service is rewarded with a tip, or gratuity, of at least 15%."

TABLE A.I—*Continued*

Sess.	ID	Message	Class	A	B
1	4	I have to do laundry tonight and I really don't want to do it! But I don't have any clean underwear left and I don't want to go commando tommorrow. We'll see what I decide tonight. This man acts funny doesn't he? But he seems cool, he's quite a character. All this mystery is kinda cool.	E	Out	R
1	5	If you will choose "In", I will choose to roll. This way, we both have an opportunity to make more than $5! ☺	P	In	R
1	6		N	Out	R
1	7	If I roll a 2–6 (you'll know when you receive the $, you will give $5.00 to a stranger. [[[then there is a line, under which is written "Sign here if you are so kind]]] Thanks. You'll still be gaining more than if I had chosen Don't roll.	P	In	R
1	8	The fairest thing to do is if you opt "IN". Then I will proceed to choose "roll." That way you and I have 5/6 chances to make money for the both of us. That's much better than just making $5 each. Increases both our chances. Thanks.	P	In	R
1	9	Choose In and I will Roll You have my word	P	In	DR
1	10	**Good luck** I do not know what I'm going to do, so I have no hints on how to advise you on choosing "in" or "out." Though it would be beneficial for me to pick don't roll and hope you pick "in", I also like to give you a chance to gain some cash. Who knows?	E	In	R
1	11	What's up? Good luck on your decision. Choose whatever. If you choose "out," you get only $10 total. If you choose "In," you can get $17 total instead of only $10. 7 bucks is a lot of money!	E	In	DR
1	12	Hey. OK I think that the best way for both of us to make a profit is for you to choose IN and for me to roll. That way we both make some money. There's no point in me not rolling because that would give you and me less profit. So I'm a roller if you're in ☺.	P	In	DR
1	13	take a risk	E	In	R
1	14	If you choose IN the first round and then I will choose Don't Roll at first. I will get $14 but then after that I will choose roll each time after the 1st role. Chances are most likely you will get $12 and I will get only $10. I will the only take 7 rolls for you to get even with me. That way we both leave with a good amount of money. Hope you have a great evening and that this works out for both of us. ☺	E	Out	R
1	15	If you choose in I'll roll. Why? If you choose out, we walk out with $10 each. If you choose IN & I choose IN then both of us coin. So it's a compromise. By agreeing to this I guarantee myself more $ than risking you choose out. So if you choose out I get $10 ($5 diff.) if you choose in I get $15 vs. $19 ($4 diff.). that's why	P	In	R

Continues

TABLE A.I—*Continued*

Sess.	ID	Message	Class	A	B
1	16		N	Out	DR
1	17		N	Out	R
1	18	Choose "In" so we can both make some $$ What are the chances me rolling a 1? I'll try my best.	P	In	R
2	1	I'm going to roll.	P	Out	R
2	2	I'll choose roll.	P	In	R
2	3	I will choose roll.	P	In	DR
2	4	I'm going to choose roll	P	In	R
2	5	choose in, & I'll roll.	P	In	R
2	6	You can have the 2 extra dollars. I'll be nice and choose to roll. ☺	P	In	R
2	7		N	In	R
2	8	Hey, choose in and I will roll. You have to like your odds that I will roll a 2,3,4,5, or 6. 5/6 odds ain't bad.	P	In	DR
2	9	If you choose "In", I'll choose Roll and you've got a 5/6 chance of getting $12.	P	In	R
2	10	Stay IN, I really need the money.	E	In	R
2	11	If you choose IN, and I roll, the chances of our getting the most $ are very high. The likelyhood of my rolling a 1 is small compared to the chances of rolling a 2–6. So we both get cash.	E	Out	DR
2	12	Hi, well I'm going to Roll so you have at least a shot for more money. I hope it works out.	P	In	R
3	1	Hopefully I'll make a lucky role.	E	Out	DR
3	2	It's much more likely that I'll roll a 2–6 and thus get more money then if we don't roll or choose out. I promise that I won't cheat you and that I'll choose to roll. ☺	P	In	R
3	3	Tee hee, this is kinda Twilight Zone – ism; Why not "go for it", eh? I hope you have a lovely evening as well.	E	Out	R
3	4	Hello fair stranger, anonymous partner . . . Choose whatever you want. Far be it from me to influence your decision, but I think you should choose "in" and I should choose "roll" and we should take the chance at both earning as much as we can. 5 chances out of 6 say it'll work, and I'm totally broke, looking to rake in stray cash however I can. I feel the luck in the air. I don't really have much else to say. Hope you're doing well, whoever you are. Yes. That's all. Random note from random human	E	In	R
3	5	Both of 'us' can earn.	E	In	DR

Continues

TABLE A.I—*Continued*

Sess.	ID	Message	Class	A	B
3	6	Ok. You're probably thinking, lets chose out, and I'll at least get 5 bucks. But... ...Chose 'IN', and I WILL chose to roll. The probability that I will roll a 2,3,4,5, or 6 is pretty high, and I think worthy of trying for. (I have no way of assuring you that I will roll ... but, its probably worth going for, you'll get $12 for finding out, where I could get $10.) x. I WILL ROLL	P	In	R
3	7	I <u>will roll</u>, so if you stay in, you've got a 5/6 chance of getting $12. If you don't mind the risk, if you stay in we'll both probably get more than $5 ... Pretty cool to get money, eh? I'm kinda bored. Hope you've had a great day so far! My country Tis of Thee Sweet Land of Liberty Of Thee I sing. Land where my fathers died Land of the Pilgrim's Pride On every mountainside Let freedom ring. George W. Bush wants you to go in! Bin Laden says "out"! ☺	P	In	R
3	8	Lets together get the most $ out of this that we can. ⇒ you 12 0 0 5 me 10 10 14 5 I promise not to do this one. ⇑ I promise I will choose to roll. You can have the extra $2 bucks. It's good karma. Thanks. I will choose ROLL in any case considering I will get the same amount no matter what you choose, as long as you choose IN. please excuse the awful handwriting. I'm trying	P	In	DR
3	9	I'm choosing ROLL, which gives you a chance to get $12 instead of $5, so stay. It's a risk, but you could end up getting a lot more.	P	In	R
3	10	If you choose in then I'm going to choose roll. This gives you a 5/6 chance of getting 12 dollars. That is 7 more than if you choose out. Since the money is free anyway – why not believe me. I'm don't lie – I promise I will choose roll.	P	In	R
3	11	If you choose <u>IN</u> you have the best opportunity to make the most money. You have a 5/7 chance of making more money! So <u>IN</u> would be your best bet. Cheers. ☺	E	In	DR
3	12	Choose IN. I promise I'll ROLL.	P	Out	R

REFERENCES

BACHARACH, M., G. GUERRA, AND D. ZIZZO (2001): "Is Trust Self-Fulfilling? An Experimental Study," Mimeo, Oxford University, Department of Economics. [1585]

BATTIGALLI, P., AND M. DUFWENBERG (2005): "Dynamic Psychological Games," Mimeo, IGIER, Bocconi University. [1592]

BAUMEISTER, R., A. STILLWELL, AND T. HEATHERTON (1994): "Guilt: An Interpersonal Approach," *Psychological Bulletin*, 115, 243–267. [1584]

——— (1995): "Personal Narratives About Guilt: Role in Action Control and Interpersonal Relationships," *Basic and Applied Social Psychology*, 17, 173–198. [1584]

BERG, J., J. DICKHAUT, AND K. MCCABE (1995): "Trust, Reciprocity, and Social History," *Games and Economic Behavior*, 10, 122–142. [1580]

BERNHEIM, D. (1994): "A Theory of Conformity," *Journal of Political Economy*, 102, 841–877. [1596]

BOLTON, G., AND A. OCKENFELS (2000): "ERC: A Theory of Equity, Reciprocity, and Competition," *American Economic Review*, 90, 166–193. [1592]

BOLTON, P., AND M. DEWATRIPONT (2005): *Contract Theory*. Cambridge, MA: MIT Press. [1579]

BRUNSON, D. (2002): *Super System: A Course in Power Poker* (Third Ed.). New York: Cardoza. [1593]

CHARNESS, G., AND M. DUFWENBERG (2006): "Supplement to 'Promises and Partnership'," *Econometrica Supplementary Material*, 74, http://www.econometricsociety.org/ecta/supmat/5008instructions.pdf. [1585,1590,1596]

CHARNESS, G., AND M. RABIN (2002): "Understanding Social Preferences with Simple Tests," *Quarterly Journal of Economics*, 117, 817–869. [1592,1593]

——— (2005): "Expressed Preferences and Behavior in Experimental Games," *Games and Economic Behavior*, 53, 151–169. [1585]

CRAWFORD, V. (1998): "A Survey of Experiments on Communication via Cheap Talk," *Journal of Economic Theory*, 78, 286–298. [1584]

CROSON, R., AND M. MILLER (2004): "Explaining the Relationship Between Action and Beliefs: Projection vs. Reaction," Mimeo, The Wharton School. [1594]

DAWES, R., J. MCTAVISH, AND H. SHAKLEE (1977): "Behavior, Communication, and Assumptions about Other People's Behavior in a Commons Dilemma Situation," *Journal of Personality and Social Psychology*, 35, 1–11. [1582]

DUFWENBERG, M. (2002): "Marital Investment, Time Consistency, and Emotions," *Journal of Economic Behavior and Organization*, 48, 57–69. [1585]

DUFWENBERG, M., S. GÄCHTER, AND H. HENNIG-SCHMIDT (2005): "The Framing of Games and the Psychology of Strategic Choice," Mimeo, University of Arizona, Department of Economics. [1595]

DUFWENBERG, M., AND U. GNEEZY (2000): "Measuring Beliefs in an Experimental Lost Wallet Game," *Games and Economic Behavior*, 30, 163–182. [1580,1585]

DUFWENBERG, M., AND G. KIRCHSTEIGER (2004): "A Theory of Sequential Reciprocity," *Games and Economic Behavior*, 47, 268–298. [1593]

DUFWENBERG, M., AND M. LUNDHOLM (2000): "Social Norms and Moral Hazard," *Economic Journal*, 111, 506–525. [1596]

ELLINGSEN, T., AND M. JOHANNESSON (2004): "Promises, Threats and Fairness," *Economic Journal*, 114, 397–420. [1582,1587,1593]

ELSTER, J. (1989): "Social Norms and Economic Theory," *Journal of Economic Perspectives*, 3, 99–117. [1595]

FARRELL, J., AND M. RABIN (1996): "Cheap Talk," *Journal of Economic Perspectives*, 10, 103–118. [1584]

FEHR, E., AND S. GÄCHTER (2000): "Fairness and Retaliation: The Economics of Reciprocity," *Journal of Economic Perspectives*, 14, 159–181. [1592]

FEHR, E., AND K. SCHMIDT (1999): "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*, 114, 817–868. [1592]

——— (2002): "Theories of Fairness and Reciprocity—Evidence and Economic Applications," in *Advances in Economics and Econometrics—8th World Congress*, ed. by M. Dewatripont, L. Hansen, and S. Turnovsky. Cambridge, U.K.: Cambridge University Press. [1592]

FISCHBACHER, U., S. GÄCHTER, AND E. FEHR (2001): "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment," *Economics Letters*, 71, 397–404. [1594]

FREY, B., AND S. MEIER (2004): "Social Comparisons and Pro-Social Behavior—Testing 'Conditional Cooperation' in a Field Experiment," *American Economic Review*, 94, 1717–1722. [1594]

FUDENBERG, D., AND D. K. LEVINE (1998): *The Theory of Learning in Games*. Cambridge, MA: MIT Press. [1592]

GEANAKOPLOS, J., D. PEARCE, AND E. STACCHETTI (1989): "Psychological Games and Sequential Rationality," *Games and Economic Behavior*, 1, 60–79. [1579,1592]

GLAESER, E., D. LAIBSON, J. SCHEINKMAN, AND C. SOUTTER (2000): "Measuring Trust," *Quarterly Journal of Economics*, 115, 811–846. [1591]

GLASNAPP, D., AND J. POGGIO (1985): *Essentials of Statistical Analysis for the Behavioral Sciences*. Columbus, OH: Merrill. [1588]

GNEEZY, U. (2005): "Deception: The Role of Consequences," *American Economic Review*, 95, 384–394. [1593]

GUL, F., AND W. PESENDORFER (2005): "The Canonical Type Space for Interdependent Preferences," Mimeo, Princeton University, Department of Economics. [1593]

GÜTH, W., P. OCKENFELS, AND M. WENDEL (1994): "Efficiency by Trust in Fairness? Multiperiod Ultimatum Bargaining Experiments with an Increasing Cake," *International Journal of Game Theory*, 22, 51–73. [1580]

HANNAN, L., J. KAGEL, AND D. MOSER (2002): "Partial Gift Exchange in an Experimental Labor Market: Impact of Subject Population Differences, Productivity Differences, and Effort Requests on Behavior," *Journal of Labor Economics*, 20, 923–951. [1585]

HUANG, P., AND H.-M. WU (1994): "More Order Without More Law: A Theory of Social Norms and Organizational Cultures," *Journal of Law, Economics and Organization*, 10, 390–406. [1584]

KERR, N. L., AND C. M. KAUFMAN-GILLILAND (1994): "Communication, Commitment, and Cooperation in Social Dilemmas," *Journal of Personality and Social Psychology*, 66, 513–529. [1587]

LEVINE, D. K. (1998): "Modeling Altruism and Spitefulness in Experiments," *Review of Economic Dynamics*, 1, 593–622. [1593]

NASH, J. (1950): "The Bargaining Problem," *Econometrica*, 18, 155–162. [1581]

RABIN, M. (1993): "Incorporating Fairness into Game Theory and Economics," *American Economic Review*, 83, 1281–1302. [1593]

ROSS, L., D. GREENE, AND P. HOUSE (1977): "The False Consensus Effect: An Egocentric Bias in Social Perception and Attribution Processes," *Journal of Experimental Social Psychology*, 13, 279–301. [1594]

ROTH, A. (1995): "Bargaining Experiments," in the *Handbook of Experimental Economics*, ed. by J. Kagel and A. Roth. Princeton, NJ: Princeton University Press, 253–348. [1582]

SELTEN, R. (1967): "Die Strategiemethode zur Erforschung des Eingeschränkt Rationalen Verhaltens im Rahmen eines Oligopolexperiments," in *Beiträge zur Experimentellen Wirtschaftsforschung*, ed. by H. Sauermann. Tubingen: Mohr, 136–168. [1585]

SIEGEL, S., AND N. J. CASTELLAN (1988): *Nonparametric Statistics for the Behavioral Sciences*. Boston: McGraw-Hill. [1589]

SOBEL, J. (2005): "Interdependent Preferences and Reciprocity," *Journal of Economic Literature*, 43, 392–436. [1592]

TANGNEY, J. (1995): "Recent Advances in the Empirical-Study of Shame and Guilt," *American Behavioral Science*, 38, 1132–1145. [1583]