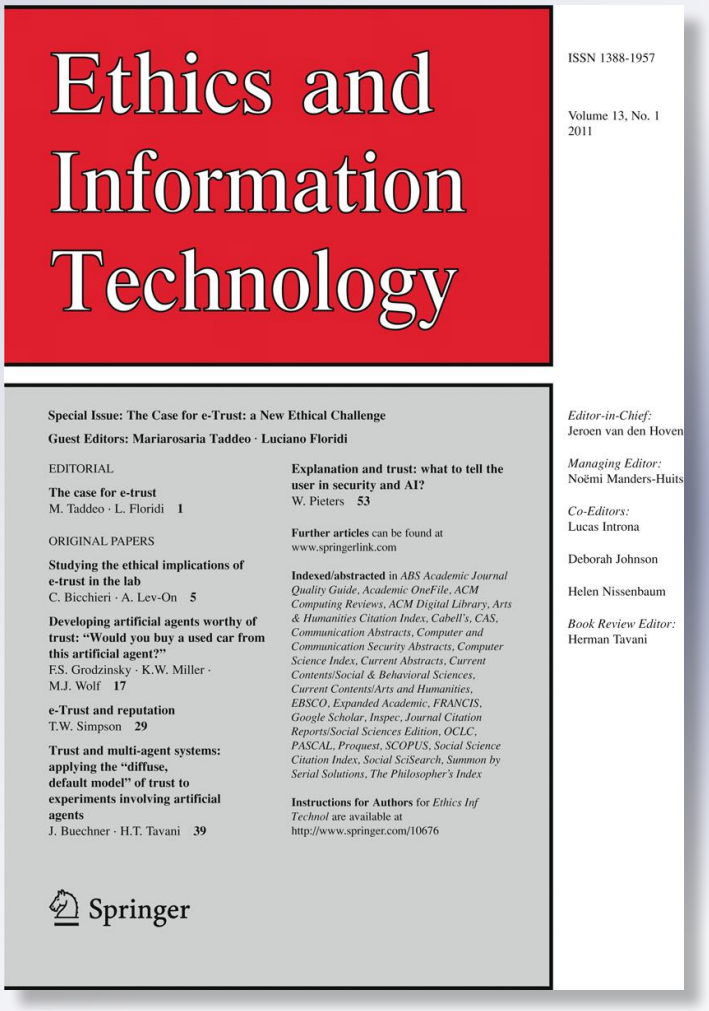
Studying the ethical implications of e-trust in the lab

Ethics and Information Technology

ISSN 1388-1957
Volume 13
Number 1

Ethics Inf Technol (2010) 13:5-15
DOI 10.1007/s10676-010-9258-y



Springer

Ethics Inf Technol (2011) 13:5–15
DOI 10.1007/s10676-010-9258-y

ORIGINAL PAPER

Studying the ethical implications of e-trust in the lab

Cristina Bicchieri · Azi Lev-On

Published online: 22 December 2010

Abstract The paper presents results of recent laboratory experiments that study if and how computer-mediated communication affects cooperation and trust. It is argued that communication medium does not matter much for trust-building and maintenance, whereas relevant pre-play communication and group size can have a major influence. The implications of the findings for the design of sites that depend on trusting communities are discussed.

Keywords Trust · Promise-keeping · Social norms · e-trust · Communication · Trust games

Communication in the lab

In environments characterized by deep hostility and distrust between populations, and especially where the barriers to exchange are not only emotional but geographical and physical as well, many hope that computer-mediated communication (specifically the Internet) may generate exposure to others' culture, beliefs and opinions, expand people's horizons and develop mutual understanding.

The experimental laboratory is a uniquely apt environment to study trust, reciprocity and cooperation, as it allows to manipulate and control for multiple variables, desegregate their effects, and conduct multiple treatments to rule out competing hypotheses. In this article we focus on social dilemma and trust games, as such games are frequently used to study the determinants of trust and cooperation in strategic interactions.

Social dilemmas are choice situations involving interdependent agents, where the choices of each influence the welfare of all. These choice situations are 'dilemmas' because the strategic setting is such that short-term rational decisions of narrowly self-interested agents lead to socially sub-optimal outcomes, i.e. the Nash equilibrium is also Pareto sub-optimal. In typical social dilemma experiments, subjects are divided into groups of size greater than two. All subjects receive an endowment and then decide to send some, all, or none of this amount to a 'group account'. The amount the subjects do not send is theirs to keep. Then, the amount accumulated in the 'group account' is multiplied by the experimenters and is equally divided among all group members. These games use the mixed-motive structure of a social dilemma, where it is individually best for subjects to keep their money in their personal account, but all are better off if everyone makes a cooperative decision and contribute their endowments to the group account.

Close cousins of social dilemma games are trust games, which are in essence sequential dilemma games. In a typical trust experiment, subjects are assigned to one of two roles: first-movers and second-movers. Experiments contain two decision periods. In the first decision period, each first-mover receives an endowment and then decides to

The article is based on previous theoretic and experimental work by Bicchieri (2006), Bicchieri and Lev-On (2007), Bicchieri, Lev-On and Chavez (2010), Lev-On, Chavez and Bicchieri (2010), and Lev-On (2009b).

C. Bicchieri (✉)
University of Pennsylvania, Philadelphia, PA, USA
e-mail: cb36@sas.upenn.edu

A. Lev-On
Institute for Research in the Social Sciences,
Stanford University, Stanford, CA, USA

A. Lev-On
School of Communication, Ariel University Center, Ariel, Israel

send some, all, or none of it to the second-mover. The amount the first-mover does not send is hers to keep. In the second decision period, the amount first-movers sent to second-movers is multiplied by the experimenters. Then, the second-mover can send some, all, or none of this amount to the first-mover. The amount the second-mover does not send is hers to keep.

In one-shot social dilemma and trust games, it might be argued that pre-play communication, and in particular promises exchanged by subjects, are "cheap talk" and therefore should not be expected to be kept. But a robust finding in the experimental literature is the positive effect of (interactive and unrestricted) communication on cooperation, which elsewhere we denote as a 'communication effect' (Bicchieri 2002; Bicchieri and Lev-On 2007; see also Ostrom 1998). Ledyard (1995), in an extensive survey of the experimental literature on public goods, singles out communication and the marginal per capita return as the two variables most conducive to cooperation. Sally (1995), in a meta-analysis of 35 years of social dilemma experiments, shows that the ability to communicate increases cooperation over base rates by 40%.

Let us now highlight several key dimensions of the communication effect:

1. As long as communication persists, cooperation rates are high and stable (Frohlich and Oppenheimer 1998; Ostrom and Walker 1991; Schmitt et al. 2000; Kinukawa et al. 2000).
2. A standard finding in iterated social dilemma experiments is that without communication cooperation gradually declines (i.e. Isaac et al. 1985; Isaac and Walker (1988); Kiesler et al. (1996); Gächter and Fehr (1999)). But cooperation rates peak after communication, *even when it takes place after a few rounds of declining cooperation* (Ostrom et al. 1992; Zheng et al. 2002; Isaac and Walker 1988). In fact, communication not only improves cooperation in the round immediately following it, but its effect *carries over* to a number of subsequent iterations (Isaac and Walker 1988).
3. When there are two separate groups, the *carryover effect* extends to the out-group. Orbell et al. (1988) allowed agents to communicate and decide which strategies to adopt for contributing to the production of a public good. After the discussion, they informed the subjects that their contributions would indeed be used to provide a public good, but a public good that only the members of another group would enjoy. Despite the unexpected change of beneficiary, 59% of the subjects gave to the out-group after discussion, significantly more than the 30% contribution rate in the control, no-communication condition (but still less than the 79% contribution rate obtained when communication was allowed and the contributions went to the original in-group beneficiaries).

A few studies demonstrate that the communication effect exists in *trust games* as well. Charness and Dufwenberg (2006) allowed unrestricted written messages from second-movers to first-movers before trust games. They found significant differences between the communication and the no-communication conditions in terms of both trusting and reciprocating behaviors.[1]

Communication and pro-social norms

What can the "communication effect" be attributed to? Dawes (1980) identified three elements of face-to face (FtF) communication that make cooperation possible: *identification*, *discussion*, and *commitment*. Experimental results enable to rule out the first two as primary causes of the communication effect, and suggest that the communication effect is caused neither by the ability to identify and 'humanize' other agents, nor by the content and dynamics of *generic* discussion (Bicchieri 2002; Bicchieri and Lev-On 2007). Communication is highly effective only when participants can discuss the game, and collectively decide how to act. In that context, participants typically make promises (to cooperate, or reciprocate), and they keep them when the time to act comes. What remains to be explained is why a promise made in a one-shot, anonymous game is taken so seriously that cooperation occurs even in the absence of any sanctioning mechanism. Or why, in repeated games, promises carry over not just to subsequent rounds (when no new discussion occurred), but even to agents who were not the original recipients of the pledge and did not promise anything in return.

The communication effect has been explained by Bicchieri (2006) in terms of her theory of social norms, which is the theoretical underpinning for our experimental research program as well. By 'social norms' we refer to informal behavioral rules that are not supported by formal sanctions. Take a norm of promise keeping. For some people, keeping promises is an important personal norm that one would follow in any circumstance, irrespective of what others do. For others, the decision to keep one's promise is conditional upon expecting most other people to keep their promises as well, and upon the belief that one is expected to fulfill one's promises, too. In this second case, we say that keeping promises is a social norm, and as such its implementation is dependent upon the expectations that individuals hold. There is much experimental evidence to support the view that important pro-social norms such as

[1] See also Ben-Ner and Putterman (2006).

reciprocity, fairness or cooperation are social rather than personal, as manipulating mutual expectations causes major behavioral changes (Bicchieri and Xiao 2009; Bicchieri and Chavez 2010).[2]

This view of social norms is crucial for explaining the effects of communication on cooperation and reciprocity as well. Cooperation/reciprocation do not occur just because people focus on a relevant norm; it is also important that the right kind of expectations are present, and certain types of communication fare better than others in creating such expectations.

According to Bicchieri (2006, see also Bicchieri and Lev-On 2007), communication *about the dilemma* has a twofold effect: it *focuses* agents on pro-social norms (particularly the norm of promise-keeping), and it also generates the kind of mutual expectations that support norm-abiding behavior. Such expectations are twofold: on the one hand individuals must have *empirical expectations* about other people's conformity with the relevant norm. Since compliance with social norms is conditional, doubting that a norm is in fact followed would diminish one's willingness to follow it. On the other hand it must also be the case that individuals have *normative expectations*, i.e., they must believe that others think they ought to obey the norm in question and may even be prepared to sanction transgressions (Bicchieri 2006).[3] Especially in the case of pro-social norms, there is a tension between self-interest and what is good for the group or society. If one expects others to cooperate, there may be the temptation to defect, and the presence of normative expectations considerably weakens such temptations.

Communication, when successful, generates a normative environment that is conducive to cooperation. There are various ways in which social norms can become salient, so that agents are led to focus on them. One way is to observe other people's normative or counter-normative behavior (Schroeder et al. 1983; Pillutla and Chen 1999). Another is to be exposed to written or verbal content that 'calls to mind' a specific norm (Cialdini et al. 1991). Yet simply focusing people on a relevant norm might not be enough to generate compliance, especially when there is some ambiguity in the decision context, or individuals receive conflicting messages (Bicchieri and Xiao 2009). Interactive, direct communication among the subjects involved in the decision situation, especially when the content of such communication involves a discussion of the decision context and mutual promises, is a highly effective mechanisms for focusing people on social norms *and* inducing compliance.

The communication effect has mostly been studied in face-to-face (FtF) settings, but it is present in computer-mediated environments as well, i.e. computer-mediated communication (CMC) produces higher cooperation rates than equivalent environments in which communication is not allowed. Here are some relevant features of the computer-mediated communication effect (Bicchieri and Lev-On 2007):

1. The communication effect varies in degree according to the richness of the communication channel. For example, videoconferencing produces cooperation rates very close to face-to-face communication, whereas text-based communication produces much less cooperation. Generally, the CMC effect approximates the FtF communication effect *the closer the communication channel comes to reproducing the features of face-to-face communication.*
2. When using CMC, communication is more normatively charged than FtF communication.[4] This could be explained by the need to 'compensate' for the lack of contextual cues in computer-mediated environments (Frohlich and Oppenheimer 1998; Rocco 1998; Brosig et al. 2003).
3. Compared to FtF communication, it takes more time to establish cooperation, especially when using 'poorer' CMC channels.
4. Especially with asynchronous communication, it is more difficult to establish 'social contracts' in CMC, and even when such agreements are reached, they are violated more frequently than agreements reached using FtF communication.

Different communication contexts can thus hamper or promote focusing on the relevant norms, and the formation of expectations that are crucial in supporting norm-abiding behavior. For example, people may be more inclined to question the credibility of online promises, with detrimental consequences for cooperation. When promises are involved, the success of face-to-face communication depends on the availability of a variety of cues that allow subjects to assess mutual intentions and form expectations about each other, all of which helps in lending credibility to mutual promises. Such indicators include *visual cues* (body language, eye contact, facial expressions, and so on), *verbal cues* (tone of voice, phrasing, fluency, manner of expressing moral rhetoric, and so on), and *social cues* (status, group membership, gender, and so on). Some of these cues are usually correlated by agents with

[2] By 'pro-social norms' we denote norms that further positive social relationships.

[3] Sanctions may be positive, as when one is praised for norm-compliance, or negative, as when one is criticized, made feel guilty or ashamed, or even ostracized by the relevant group (Bicchieri 2006 Ch.1).

[4] For example, greater use of (empty) threats against potential free-riders is common in CMC environments.

Table 1 Mean trust, reciprocity, and first-mover expectations by communication relevance and medium (N = 64)

	Control (N = 32)	FtF-relevant (N = 14)	CMC-relevant (N = 14)	FtF-irrelevant (N = 18)	CMC-irrelevant (N = 18)
Trust	2.63 (0.36)	5.57 (0.46)	5.14 (0.57)	4.17 (0.49)	3.28 (0.61)
Reciprocity	1.92 (.48)	7.57 (0.96)	5.14 (1.33)	3.33 (1.05)	1.94 (0.78)
Expected reciprocity	3.54 (0.53)	8.36 (0.69)	7.43 (0.96)	5.56 (0.91)	4.28 (0.93)

trustworthiness, and their presence or absence can have important motivational consequences via the formation (or impairment) of mutual expectations of promise-keeping behavior.

The study of computer-mediated communication and the conditions under which mutual promising and collective commitments are likely to take place, is the subject of ongoing research. Some of our recent studies intended to further explore the 'communication effect' in computer-mediated settings, and refine the understanding of its determinants. So far it has been shown that the medium of communication matters; the question now shifts to what happens when the medium in use for pre-play talk interacts with additional variables.

Experiment 1: media richness and communication relevance

In the first experiment surveyed below (Bicchieri et al. 2010), two dimensions of the communication effect in trust games were manipulated: communication relevance (i.e., the situational relevance of what is communicated), and the richness of the communication medium (CMC vs. FtF). Relevant communication may matter due to the possibility to talk about the strategic situation that the participants are facing, and to make non-binding pledges about future actions.[5]

To study the interaction between media richness and relevance, 64 participants were recruited at the University of Pennsylvania. Each experimental session consisted of a sequence of three identical trust games. For each game, the first-mover had 6 USD, any dollar amount of which could be sent to the second-mover. The amount the second-mover received from the first-mover was tripled. The second-mover could then send any dollar amount back to the first-mover.[6] Before playing the game, participants were allowed a few minutes of discussion. Half of the participants were engaged in face-to-face discussion, and another half had computer-mediated chats. In roughly half of the experimental sessions, participants were involved in "relevant" communication, i.e. they were allowed to discuss any topic except those pertaining to their identities and their decisions or earnings from the previous condition, if there was one. In the remaining sessions, the instructions required that participants discuss only the following questions (adapted from Buchan et al. 2006), which were not relevant to the game: *What are the three most populated cities in the world? What are the three most populated cities in the US? How many people live in Philadelphia and the surrounding suburbs? How many counties are there in Pennsylvania?* ("irrelevant" communication).[7] In these last sessions, participants were explicitly instructed not to talk with their partner about the game they were about to play.

After making their decision, first-movers were asked about their expectations of second-movers reciprocation. We wanted to know whether expectations differed depending on the communication medium and condition, and if expectations predict the level of trust. Note that by *trust* we mean the amount of dollars sent by first-movers ($0 through $6), by *reciprocity* we mean the amount returned by second-movers, relative to the amount sent, and by *expected reciprocity* we mean the amount the first-mover expected to be returned by the second-mover, relative to the amount sent.

Table 1 (Bicchieri et al. 2010) summarizes the responses across the five combinations of communication relevance and medium. Both relevance and medium had large, positive effects on trust, reciprocity, and expected reciprocity relative to the control condition.[8] Relevant, face-to-face communication had the largest effect on all three variables, whereas relevant, computer-mediated communication had the second largest effect.

Figure 1 shows the distribution of trust across the five conditions. First-movers were most trusting in the relevant

[5] Support for this hypothesis is also found in Bouas and Komorita (1996); Dawes et al. (1977); Gächter and Fehr (1999).

[6] Participants were paired randomly with a different partner for each game, and this was common knowledge. First-movers did not receive feedback on the amount that the second-mover returned until the end of the experimental sessions.

[7] Bicchieri (2002) discusses several social dilemma experiments in which the subject of conversation was relevant for the participants (for example, a rise in college tuition), but the behavioral result was not different from a no-communication condition. The data she discussed show that only communication about the experiment have an effect on cooperation.

[8] In the control condition, no communication was allowed.

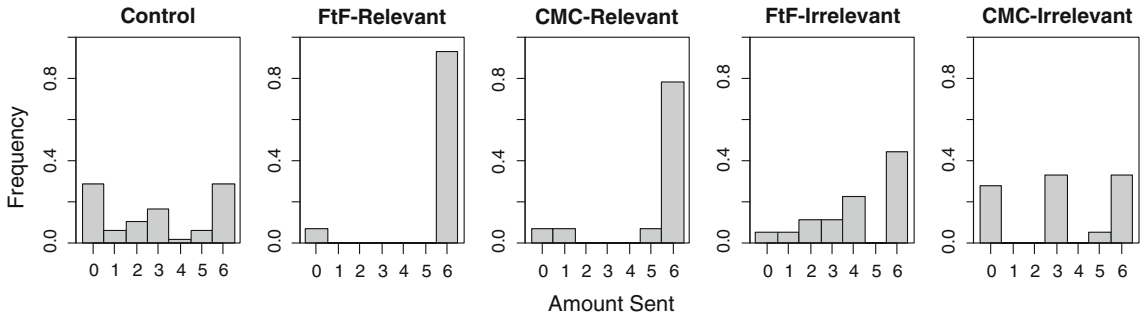


Fig. 1 Distribution of trust by communication medium and relevance

Table 2 Estimates for expected reciprocity (proportion) as a predictor of trust (N = 64)

Variable	Estimate	SE	T
Intercept	0.37	0.22	1.69
Expected percentage reciprocity	10.08	0.55	18.284****

$R^2 = 0.78$

**** $P < .0001$

communication conditions, in which the majority of participants sent their entire endowment of 6 USD.

When we conducted a simultaneous regression of trust on control, communication relevance, and medium, *F* tests revealed significant effects of control (i.e. communication vs. no-communication) ($F(1,93) = 4.47$, $P = .037$) and relevance ($F(1,93) = 8.54$, $P = .004$), *but not of medium* (CMC vs. FtF) ($F(1,93) = 1.56$, $P = .22$). Moreover, there was no interaction between communication medium and relevance ($F(1,92) = 0.01$, $P = .92$). Controlling for the other variables, first-movers had over five time greater odds of sending each dollar when communication was relevant. Relative to the other conditions, participants had three times lower odds of sending each dollar in the control condition (Bicchieri et al. 2010).

Next we wanted to determine whether expected reciprocity predicted the first-mover's level of trust.[9] The estimated coefficients in Table 2 show that trust increases with expected reciprocity. When the first-mover expected nothing to be returned, the predicted amount sent was only .36 dollars. For each percent of the amount sent that the first-mover expected to be returned, the first-mover sent an additional .10 dollars. Thus, the median expected reciprocity (proportion) of .45 resulted in a $5 increase in the amount returned. The R^2 value of 0.79 indicates that a large percentage of the variance in trust is explained by the expected percentage reciprocity.

These results suggest that the behavior of first-movers was strongly determined by their expectations of second-movers' reciprocation. The variable most conducive to creating such expectations was—paraphrasing McLuhan (McLuhan and Fiore 1967)—not the medium, but rather the message. Though the richness of the medium of communication (FtF vs. CMC) failed to produce significant differences in first-mover investments, such investments were significantly higher following unrestricted communication than restricted or no communication. Recall that unrestricted communication could include strategic discussion of the game and promise-making; according to our transcripts, all subjects who participated in the unrestricted communication were involved in both.

The medium of communication, however, had an effect on reciprocity. The pattern of second-mover return was highly bi-modal, with participants returning either nothing or exactly half of the maximum (i.e., returning 9 out of 18 USD). This pattern depended on the conditions, with almost all second-movers returning 9 USD in the FtF-relevant condition, and almost all returning nothing in the control condition. Although this pattern was partly due to the different levels of trust across conditions, it was also clear that second-movers behaved qualitatively differently across (medium of) communication conditions, even after accounting for first-movers' levels of trust. The probability of returning each available dollar increased with the amount trusted, but increased more rapidly for FtF than for CMC. Our result is consistent with earlier findings that individuals engage in positive reciprocation, especially following verbal communication.

It should be noted, however, that actual reciprocity is always lower than expected reciprocity, and this result is constant across Trust games (Camerer 2003). Yet trusting on the part of first-movers may be rational, insofar as trusting acts as a signal, whose intended effect is to focus the recipient on a reciprocity norm. If such a norm exists and is shared, then it is rational to trust insofar as one

[9] Expected reciprocity was converted into the *expected percentage reciprocity* (the amount expected to be returned divided by the amount available × 100), and then regressed trust on this variable using ordinary least squares. This conversion was necessary to control for the dependency of the maximum amount that could be returned on the amount sent.

Table 3 Mean trust, reciprocity, and first-mover expectations by communication medium and group size

	No-communication	FtF-dyadic	CMC-dyadic	FtF-group	CMC-group
Trust	3.03 (.46)	5.57 (.46)	5.14 (.57)	4.12 (.69)	3.94 (.68)
Reciprocity	1.83 (.60)	7.57 (.96)	5.14 (1.33)	3.62 (1.11)	2.12 (.95)
Expected reciprocity	3.50 (.67)	8.36 (.69)	7.43 (.96)	4.31 (1.17)	5.00 (1.14)

Parenthesized values are standard errors of the mean

believes that in so doing one will trigger reciprocation even when the material incentives to reciprocate are absent (Bicchieri et al. 2010, 2011). The fact that second-movers usually return less than expected may be explained by a theory of norm compliance as conditional upon having the right sort of expectations (Bicchieri 2006). In the anonymous environment, there is no risk of being punished, and thus the pull of the norm, though present, is less strong. However, when relevant pre-play communication is allowed, second-movers are focused on reciprocation, and promises do matter, especially when first-movers show full trust.

The data presented above only apply to dyadic communication. Yet communication often involves groups, thus it remained to be seen if the effects of communication relevance and medium hold when individuals are involved in group discussion, be it face-to-face or computer mediated.

Experiment 2: group vs. dyadic pre-play talk

In a second experiment (Lev-On, et al. 2010), the richness of the communication medium and the *size* of the communicating group were manipulated. We hypothesized that dyadic communication would be more conducive to trust and reciprocation than group communication, since in dyadic communication the players directly communicate with their counterparts in the actual game that follows, and their actions have a clear consequence for the other player—as agents' choices directly punish or reward a single identifiable person. Also, a player's promise to invest or reciprocate is directed to the person with whom they later play the game, thus triggering an additional motive—guilt aversion—for players not to break their promises, in spite of their cheap-talk status. Yet group communication may also be conducive to trust. A multiplayer pre-play conversation may encourage subjects to focus on public reasons and channel the discussion into a cooperative path, although participants are eventually paired with only a single person from the group, without knowing in advance who this person will be.

For this new experiment, 60 college students were recruited. Again, each session consisted of three sequential trust games. For each game, the first-mover had 6 USD, any dollar amount of which he or she could send to the second-mover. The amount second-movers received from the first-movers was tripled by the experimenter. The second-mover could then send any dollar amount back to the first-mover. As in the previous experiment, participants were paired randomly with a different partner for each game, and this was common knowledge. First-movers did not receive feedback on the amount that the second-mover returned until the end of the experimental sessions.

Prior to the first game, participants were not allowed to communicate, making it a no-communication/control condition. Prior to the second game, participants communicated in real time via *computer-based text chat* for five minutes with the person they were paired with or, in the group condition, for ten minutes with a group of eight people, one of which was the (anonymous) person they were paired with in the trust game.[10] Messages entered by each participant appeared in a chat window visible to all group members. Prior to the third game, participants communicated *face-to-face* for two minutes with the person with whom they were paired or, in the group condition, they had seven minutes to communicate with an eight-person group.[11] Participants then returned to their computer stations and made their decisions in the game privately.

Table 3 (Lev-On et al. 2010) summarizes the responses for the five combinations of group size and medium. Both group size and medium had large, positive effects on trust, reciprocity, and expected reciprocity relative to no communication. Dyadic, face-to-face communication had the largest effects on all three variables, whereas dyadic,

[10] The instructions specified that participants were allowed to discuss any topic except those pertaining to (1) their identities or (2) their decisions or earnings from the previous condition, if there was one. Note that we had to allow *more time* to group discussion, in order to let each participant have their say and have time to read others' messages.

[11] Again, we let group participants have more discussion time, to give each of them the possibility of expressing their opinion. Note that the times allotted to both dyadic and group FtF communication are less than the times allotted to CMC participants. This is due to the fact that exchanging computer messages takes more time than direct verbal communication.

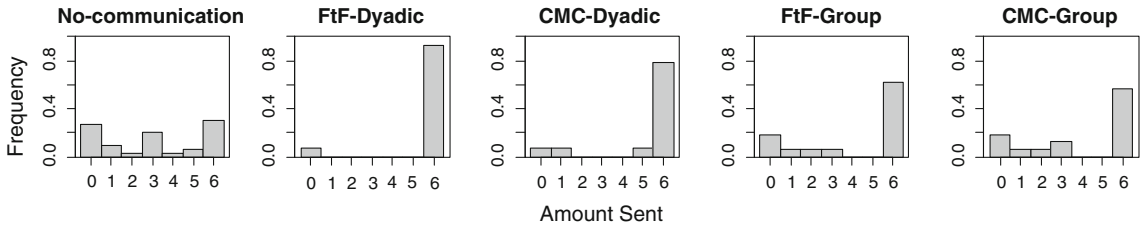


Fig. 2 Distribution of trust by group size and medium

Table 4 Estimates for expected percentage reciprocity as a predictor of trust (N = 90)

Variable	Estimate	SE	T
Intercept	1.27	0.29	4.41^{****}
Expected Percentage Reciprocity	8.75	0.73	11.94^{****}

$R^2 = 0.61$

**** $P < .0001$

computer mediated communication had the second largest effect.

Figure 2 shows the distribution of trust across the five communication conditions. To test for effects of medium and group size, we conducted a simultaneous regression of trust on the dummy variables no-communication, dyadic, and FtF. F-tests revealed significant effects of communication ($\chi^2(1) = 4.88$, $P = .03$) and group size ($\chi^2(1) = 26.56$, $P < .0001$), but not of medium ($\chi^2(1) = 1.39$, $P = .24$). Moreover, there was no interaction between medium and group size ($\chi^2(1) = 1.10$, $P = 0.29$). Thus, trust levels depended on the presence of communication and whether that communication was in dyads or groups, but again they *did not depend on whether communication was face-to-face or computer-mediated*.

Next we checked whether expected reciprocity predicted the first-mover's level of trust. The results again demonstrate (in Table 4) that trust increases with the expected percentage reciprocity. When the first-mover expected nothing to be returned, the predicted amount sent was only 1.27 dollars. For each percent of the amount sent that the first-mover expected to be returned, however, the first-mover sent an additional .88 dollars. Thus, the median expected percentage reciprocity of 50% resulted in a $4.4 increase in the amount returned. Again, the R^2 value of 0.61 indicates that a large percentage of the variance in trust is explained by the expected percentage reciprocity.

Our results show, again, that the behavior of first-movers is strongly determined by their expectations of second-movers' reciprocation. This time, the variable most conducive to creating such expectations was not the medium of communication, but rather the number of communicators. Investments in the dyadic communication conditions were significantly higher than in the group communication conditions, which were in turn significantly higher than in the no-communication condition. In the dyadic conditions, almost all first-movers sent their entire endowment, compared to only 60% in the group conditions and 30% when there was no communication.

As to the behavior of second-movers, their pattern of returns was bi-modal, again, with many participants returning nothing or exactly half of the maximum (i.e., 0 or 9 USD). This pattern depended on the communication condition. For example, almost all participants in the FtF-Dyadic condition returned 9 USD, but almost all participants in the no-communication and CMC-Group conditions returned nothing. When the first-mover sent the entire endowment, reciprocity was greatest in the dyadic conditions, with 80% of second-movers returning at least half of the amount received, compared to only 42% in the group conditions and 44% when there was no communication. Across group sizes, when first-movers sent less than their entire endowment, second-movers tended to send back little. Although this pattern was due in part to the different levels of trust across conditions, second-movers behaved qualitatively differently across conditions, even after accounting for first-movers' levels of trust. The probability of returning each available dollar increased with the amount trusted, but increased more rapidly for the dyadic conditions, and most rapidly for the FtF-Dyadic condition.

As a general rule, higher levels of trust, reciprocation, and expectations of reciprocity were recorded in the dyadic conditions, compared to the group conditions. Since a promise to trust/reciprocate was far more common in dyadic communications, this result is not surprising. In the group condition, however, when promises to trust/reciprocate were collectively made, trust and reciprocation were far more frequent than in the control, no-communication condition.

To summarize: communication richness (FtF vs. CMC) *failed* to produce significant differences in first-mover investments.[12] This time, the *size* of the communicating

[12] This finding does not conform to other experimental results (see Bochet et al. 2006; Brosig et al. 2003; Frohlich and Oppenheimer 1998; Bos et al. 2001; Rocco 1998; Zheng et al. 2002) that found significant differences in cooperation rates between 'richer' and 'poorer' communication conditions. In these studies, however, the number of communicators in a group remained constant.

group did make a difference: The amounts sent were significantly higher in the dyadic communication conditions than in the group communication and no-communication conditions. Our results suggests that, in addition to the influence of communication media found in earlier experiments, there are additional variables (such as the content of conversation and the number of discussants) that mitigate the perception of the credibility of promises and generate expectations and behavioral consequences. In a group context, unless all members promise to trust/reciprocate, it is more difficult to establish expectations conducive to support such behaviors. Furthermore, promises are much more frequent in groups engaging in face-to-face communication than in groups that communicate via computer, thus explaining the differences in trust/reciprocation we found among these two conditions.

Discussion and conclusions: the ethical implications of e-trust lab findings

The experimental results that were surveyed show that communication matters, but the "communication effect" is embedded within the normative background primed by experimenters. When only communication is manipulated (no-communication vs. communication), cooperation and trust tend to be significantly higher in environments where communication is allowed. But when additional variables are manipulated, things get more complicated. In dyadic communication, the medium of communication has almost no effect on trust, and the relevance of what is communicated has a major effect. Reciprocity, on the contrary, is affected by the richness of the communication medium. In group communication, group size and relevance of communication matter to trust, whereas the richness of the communication media did not matter. For reciprocity, again, the richness of the communication medium mattered, along with group size.

Our data show that, in addition to the influence of communication media found in earlier experiments, there are further variables (such as the content of conversation and the number of discussants) that mitigate the perception of the credibility of promises and affect expectations (with behavioral consequences). Indeed, the available data show the advantages of dyadic over group communication and of relevant over irrelevant communication, over and above the impact of the richness of the communication medium.

Communication has an impact on pro-social behavior because, among other things, it conveys information about the other party, and this information can be used to assess the credibility, trustworthiness, etc. of those we interact with. Information, however, has a darker side. It has been shown (Lev-On 2009a) that people can use information to discriminate against other groups—for example, first-movers may send more or less money depending on the ethnicity of second-movers. Thus, when *group identities are involved*, it seems that *anonymity* can actually *produce more trust*. It remains possible that limited information leads to stereotyping, focusing people, in the absence of additional trust-supporting cues, not on pro-social norms, but rather on norms of in-group favoritism that impair possible interactions with members of opposing groups. One should therefore be particularly cautious in assessing the interaction of information, communication, and the priming of social norms.

The "communication effect", particularly in computer-mediated environments, is a complicated and nuanced phenomenon. Richer communication can be conducive to cooperation, but at other times it may jeopardize trust. And quite often, the communication effect is mitigated by other variables that focus subjects on pro- (or anti-) social norms. The medium of communication does not matter in and of itself, but due to its ability to focus subjects on pro-social norms (promise-keeping, reciprocity, cooperation), and to facilitate (or hamper) the formation of mutual expectations that support norm compliance. As the experimental results show, subjects can be focused on norms in a variety of ways, and communication richness is only one among them. Communication relevance is crucial, as it allows participants to discuss their situation and make mutual promises. For example, a rich communication medium (FtF) that does not allow subjects to make promises to each other may be detrimental to trust, as our experimental results show.

But there is good news as well. Managers, organizers, moderators and software designers have a significant degree of control over online communicative and deliberative environments, which includes managing membership and content (i.e. agenda setting, facilitating and encouraging discussions, moderating discussions, preventing 'flaming' and removing inappropriate posts, archiving old threads of discussion), framing and enforcing policies regarding accepted behaviors and sanctions, as well as technical and financial management (Preece 2000). Familiarity of site managers with experimental results, such as the ones we presented, can motivate institutional design that would lead participants to focus on and adhere to pro-social norms. Cooperation and trust may indeed be facilitated, survive and be continuously adhered to in such pro-social environments.

The results we presented here have interesting online institutional design implications. For example, think of virtual teams or distributed, ad-hoc workgroups where individuals cooperate with other team members whose identity they may not know in advance. Since dyadic, and not whole-group communication, seems crucial for

accomplishing a group's goal, our experimental data show that a 'motivational' conversation with all workgroup members is not a proper substitute for direct communication, whenever possible, with the person whom one should eventually trust and with whom one should cooperate. A group's goal may be fruitfully divided into small tasks, each assigned to few members who would then communicate among themselves.

An additional implication involves the content of communication. Many online exchange sites allow some form of communication between future exchange partners. But these forms of communication vary widely. It is likely that the closer the sites come to emulating face-to-face communication, the more conducive they are to triggering mutually-beneficial exchanges. On the other hand, when sites do not allow subjects to convey to each other much information relevant to their future exchanges (for example, when they provide only information about identities but not about past behaviors), they may disable the formation of empirical and normative expectations of trust and trustworthiness, and thus jeopardize the success of computer-mediated exchange.

The experimental results we presented further support the arguments of scholars like Lessig (2006) and Flanagan et al. (2008), who claim that technological designs embed values, which in turn have cognitive and behavioral implications. Our data show that their arguments apply to the formation of e-trust as well. In our case, it is evident that the design of computerized environments, within which subjects form trust judgments, can focus subjects on certain aspects of the issue at hand, dimming or neglecting other aspects. According to the focus theory of norms, the empirical and normative expectations that are formed in such environments have significant behavioral implications (Bicchieri 2006).

The important questions to be asked are thus how much and how far we may go in manipulating the information that is shared or presented to online participants, and what design features would help build on communication effects to increase trust when it is appropriate, and to shape appropriate expectations about members' participation. The experiments we performed concentrated on a number of features that are typically present in trust-supporting environments: the number of discussants, the topics of discussion, and the information available—which are all relevant to the achievement of trust, especially among members of heterogeneous groups. Web managers and designers interested in promoting trust and cooperation may not only find the experimental information useful but, we want to argue, may frame their design choices in light of such information**,** which in turn raises interesting ethical concerns and dilemmas**.** Let us conclude by commenting on three such concerns.

The first design dilemma involves the question of *how much* information about the behavior of participants should e-trust site managers provide. Experiments show that people respond in a negative (i.e., uncooperative) way when matched with known free-riders. When information is available about behavior of other subjects who participate in the game, subjects with high normative expectations may be faced with information about low investment levels of other players. As a consequence, such players will change their empirical expectations about what other players will do, adjust their contribution accordingly by diminishing contribution rates, and set off a snowball effect of reduced or null contributions (see Kurzban et al. 2001; Wilson and Sell 1997; Isaac et al. 1985; Bicchieri and Xiao 2009).

Site designers, then, face a tough choice. Should they, in the name of transparency and openness and in the hope of luring additional contributors, make past contributions visible and information transparent, and provide indicators about members' contributions and production levels during the production of continuous public goods? Should they take the risk-averse path and hide information, in order to avoid betraying the expectations of contributors, if indeed the going gets rough? Or should they select a middle ground, for example by selectively allowing some pieces of information to expire, or by presenting or weighing them differently? A utilitarian view would suggest that, if the probability of enhancing contributions is sufficiently high, it makes sense to conceal potentially damaging information. On the other hand, this policy violates participants' reasonable expectation to know past contribution patterns.

A second design concern has to do with *how* to present information. To understand the importance of this point, let us look at reputation systems—that is, systems established to overcome trust problems by gathering inputs from large numbers of exchange partners (Lev-On 2009b). Resnick et al. (2000) argue that the way feedback and ratings are displayed in such systems can have tremendous impact over the exchange decisions of participants, as they focus the attention on certain aspects of the potential partners that may have been otherwise overlooked. Again, these design decisions can push subjects toward, or away from, trust and cooperation. We may ask whether it is fair to manipulate information to the effect of inducing trust and cooperation, and whether the final result justifies such 'hidden' control.

A simplistic view of trust formation would suggest that, the more visible to each other the participants are, the more likely it is that they will come to trust each other. Several online methods may be used to enhance the public visibility of participants, such as queues, discussion threads, and sophisticated visualization techniques. Disseminating personal information about participants, however, raises further ethical concerns. Let us focus on a simple method to enhance visibility: showing pictures of participants. Experiments

show that physical appearance matters to trust; for example, Eckel (2007a, b) found a 'beauty premium' in an online trust game that involved participants' pictures. Wilson and Eckel (2006), in another game involving pictures, found both a 'beauty premium' and a weak positive effect of smiling on monetary transfers. Such findings pose a dilemma to website managers that care about facilitating trust among members: Should they allow users to upload photos? Since individuals may base their trust judgments at least in part on cues derived from the pictures, and associate such cues with trustworthiness, they may come to trust or refrain from trusting based on appearances, which may lead to unforeseen consequences. It is interesting to note that a similar problem may exist in e-voting systems that allow candidates to include their pictures alongside their names, and enable voters to get a glimpse of the candidates *as they vote*. Such institutional designs may enable attractive political candidates to 'monetize' their 'beauty premium'—a factor that may have been neutralized to some extent in earlier and 'poorer' voting technologies.

In both cases, a genuine attempt to decrease social distance, by means of pictures or other information, may lead people to easily trust someone they should not reasonably trust or, conversely, may impair trust formation when in fact the other person is perfectly trustworthy. Knowing people's cognitive biases may thus lead to a form of 'soft paternalism', where hiding some information, manipulating it, or even preventing its diffusion may be justified by the greater good (in term of social cooperation) such actions are expected to produce.

Our aim here is not to support particular design or policy choices. Rather, we want to highlight how familiarity with experimental results such as those we have described can raise key ethical questions for those who manage and design online trust-supporting environments. As we have seen, such environments introduce a variety of practical and ethical concerns that may have far-reaching consequences for trust and cooperation. IT managers and designers have a significant control over the environments in which trust judgments are made. As their choices have ethical implications, awareness of experimental results on the determinants of trust and cooperation can serve not only as guide for usability and sociability (Preece 2000), but as an ethical compass as well.

Acknowledgment Authors wish to thank three anonymous referees for several useful comments and insights.

References

Ben-Ner, A. & Putterman, L. (2006). Trust, communication and contracts: Experimental evidence.' Working Papers 2006-23, Brown University, Department of Economics.

Bicchieri, C. (2002). Covenants without swords: Group identity, norms, and communication in social dilemmas. *Rationality and Society, 14*(2), 192–228.

Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms*. New York: Cambridge University Press.

Bicchieri, C., & Chavez, A. (2010). Behaving as expected: Public information and fairness norms. *Journal of Behavioral Decision Making, 23*(2), 161–178.

Bicchieri, C., A. Lev-On & Chavez, A. (2010). 'The medium or the message? Communication richness and relevance in trust games.' *Synthese, 176*(1), 125–147.

Bicchieri, C., & Lev-On, A. (2007). Computer-Mediated communication and cooperation in social dilemmas: An experimental analysis. *Politics, Philosophy and Economics, 6*, 139–168.

Bicchieri, C., & Xiao, E. (2009). Do the right thing: But only if others do so. *Journal of Behavioral Decision Making, 21*, 1–18.

Bicchieri, C., Xiao, E., & Muldoon, R. (2011). Trustworthiness is a social norm, but trusting is not. *Politics, Philosophy and Economics* (forthcoming).

Bochet, O., Page, T., & Putterman, L. (2006). Communication and punishment in voluntary contribution experiments. *Journal of Economic Behavior and Organization, 60*, 11–26.

Bos, N., Gergle, D., Olson, J. S., & Olson, G. M. (2001). *Being there versus seeing there: Trust via video. Proceedings of CHI 2001: 291–2*. New York: ACM Press.

Bouas, K. S., & Komorita, S. S. (1996). Group discussion and cooperation in social dilemmas. *Personality and Social Psychology Bulletin, 22*, 1144–1150.

Brosig, J., Ockenfels, A., & Weimann, J. (2003). The effect of communication media on cooperation. *German Economic Review, 4*, 217–241.

Buchan, N. R., Croson, R. T. A., & Johnson, E. J. (2006). Let's get personal: An international examination of the influence of communication, culture, and social distance on other regarding preferences. *Journal of Economic Behavior and Organization, 60*, 373–398.

Camerer, C. (2003). *Behavioral game theory: Experiments on strategic interaction*. NJ, Princeton: Princeton University Press.

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

Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology, v. 24* (pp. 201–234). San Diego, CA: Academic Press.

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology, 31*, 169–193.

Dawes, R. M., McTavish, J., & Shaklee, H. (1977). Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. *Journal of Personality and Social Psychology, 35*, 1–11.

Eckel, C. C. (2007a). People playing games: The human face of game theory (presidential address). *Southern Economic Journal, 73*(4), 841–857.

Eckel, C. C. (2007b). People playing games: The human face of game theory. *Southern Economic Journal, 73*(4), 841–857.

Flanagan, M., Howe, D., & Nissenbaum, H. (2008). Embodying values in technology: Theory and practice. In J. van den Hoven & J. Weckert (Eds.), *Information technology and moral philosophy*. Cambridge University Press: Cambridge.

Frohlich, N., & Oppenheimer, J. (1998). Some consequences of e-mail vs. face to face communication in experiment. *Journal of Economic Behavior and Organization, 35*, 389–403.

Gächter, S., & Fehr, E. (1999). Collective action as a social exchange. *Journal of Economic Behavior and Organization, 39*, 341–369.

Isaac, R. M., & Walker, J. M. (1988). Communication and free-riding behavior: The voluntary contribution mechanism. *Economic Inquiry, 26*, 585–608.

Isaac, R. M., McCue, K., & Plott, C. R. (1985). Public goods provision in an experimental environment. *Journal of Public Economics, 26*, 51–74.

Kiesler, S., Sproull, L., & Waters, K. (1996). A prisoner's dilemma experiment on cooperation with people and human-like computers. *Journal of Personality and Social Psychology, 70*, 47–65.

Kinukawa, S., Saijo, T., & Une, M. (2000). Partial communication in a voluntary-contribution-mechanism experiment. *Pacific Economic Review, 5*, 411–428.

Kurzban, R., McCabe, K., Smith, V. L., & Wilson, B. J. (2001). Incremental commitment and reciprocity in a real time public goods game. *Personality and Social Psychology Bulletin, 27*(12), 1662–1673.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. H. Kagel & A. E. Roth (Eds.), *Handbook of experimental economics* (pp. 111–194). Princeton: Princeton University Press.

Lessig, L. (2006). *Code, version 2.0*. New York: Basic Books.

Lev-On, A. (2009a). The virtues of anonymity: Evidence from trust games. Manuscript.

Lev-On, A. (2009b). Cooperation with, without trust online. In K. S. Cook, C. Snijders, V. Buskens, & C. Cheshire (Eds.), *ETrust: Forming relationships in the online world* (pp. 292–318). New York: Russell Sage.

Lev-On, A., Chavez, A., & Bicchieri, C. (2010). Group and dyadic communication in trust games. *Rationality and Society, 22*(1), 37–54.

McLuhan, M., & Fiore, Q. (1967). *The medium is the message*. New York: Bantam Books.

Orbell, J. M., van de Kragt, A. J., & Dawes, R. M. (1988). Explaining discussion-induced cooperation. *Journal of Personality and Social Psychology, 54*, 811–819.

Ostrom, E. (1998). A behavioral approach to the rational choice theory of collective action- presidential address of the American political science association 1997. *American Political Science Review, 92*, 1–22.

Ostrom, E., & Walker, J. M. (1991). Communication in a commons: Cooperation without external enforcement. In T. R. Palfrey (Ed.), *Laboratory research in political economy* (pp. 287–322). Ann Arbor: University of Michigan Press.

Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *American Political Science Review, 86*, 404–417.

Pillutla, M. M., & Chen, X. (1999). Social norms and cooperation in social dilemmas: The effects of context and feedback. *Organizational Behavior and Human Decision Processes, 78*, 81–103.

Preece, J. (2000). *Online communities: Designing usability and supporting sociability*. New York: John Wiley.

Resnick, P., Zeckhauser, R., Friedman, E., & Kuwabara, K. (2000). Reputation systems. *Communiations of the ACM, 43*, 45–48.

Rocco, E. (1998). *'Trust breaks down in electronic contexts but can be repaired by some initial face-to-face contact'. Proceedings of CHI 1998: 496–502*. New York: ACM Press.

Sally, D. (1995). Conversation and cooperation in social dilemmas. *Rationality and Society, 7*, 58–92.

Schmitt, P., Swope, K., & Walker, J. (2000). Collective action with incomplete commitment: Experimental evidence. *Southern Economic Journal, 66*, 829–854.

Schroeder, D. A., Jensen, T. D., Reed, A. J., Sullivan, D. K., & Schwab, M. (1983). The actions of others as determinants of behavior in social trap situations. *Journal of Experimental Social Psychology, 19*, 522–539.

Wilson, R. K., & Eckel, C. C. (2006). Judging a book by its cover: Beauty and expectations in a trust game. *Political Research Quarterly, 59*(2), 189–202.

Wilson, R. K., & Sell, J. (1997). Liar, liar… cheap talk and reputation in repeated public goods settings. *Journal of Conflict Resolution, 41*, 695–717.

Zheng, J., Veinott, E., Bos, N., Olson, J. S., & Olson, G. M. (2002). *Trust without touch: Jumpstarting trust with initial social activities. Proceedings of CHI 2002: 141–6*. New York: ACM Press.