Politics, Philosophy & Economics
http://ppe.sagepub.com

Computer-mediated communication and cooperation in social dilemmas: an experimental analysis

Cristina Bicchieri and Azi Lev-On

Politics Philosophy Economics 2007; 6; 139
DOI: 10.1177/1470594X07077267

The online version of this article can be found at:
http://ppe.sagepub.com/cgi/content/abstract/6/2/139

Published by:
SAGE Publications
http://www.sagepublications.com

On behalf of:
The Murphy Institute of Political Economy

1. Characteristics of the communication effect

One robust experimental finding in the study of social dilemmas is the positive effect of communication on cooperation, which we denote as the 'communication effect'. The communication effect has been studied in many experimental settings in which communication among subjects was allowed before or during social dilemma games.[2] John Ledyard, 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. Likewise, David Sally, in a meta-analysis of 35 years of social dilemma experiments, shows that the ability to communicate increases cooperation over base rates by 40 percent.[3]

Note that in the majority of the experiments reported here, communication took the form of *unrestricted* face-to-face (FtF) discussion ranging from two to ten minutes. Unless otherwise stated, subjects were in the same location, could see each other, and were allowed to talk about the game or any other issue. When they discussed the game, subjects could make nonbinding promises to their group. Some experiments we shall later discuss modified and manipulated the communication conditions in order to isolate those elements of the communication process that affect cooperation.[4] Table A1 in the Appendix specifies the communication conditions in the experiments we consider.

To appreciate the extent of the communication effect, it is useful to consider some of the most interesting experimental results. In an early study, Dawes, McTavish, and Shaklee found that in a commons dilemma experiment cooperation in the FtF communication condition reached 72 percent, significantly more than the 31 percent cooperation rate of the no-communication condition.[5] Ostrom and Walker similarly report that in a series of experiments continuous FtF communication more than tripled cooperation in a repeated commons game; in fact, cooperation rates increased from 30 percent before communication to 98 percent after communication.[6] Frohlich and Oppenheimer also found that FtF communication tripled cooperation in a public-good game.[7] Both Bornstein, in variants of public-good games, and Hackett, Schlager, and Walker, in a commons dilemma experiment, show that FtF communication and promising almost entirely eliminated defections.[8]

Let us now specify several key dimensions of the communication effect:[9]

1. As long as communication persists, cooperation rates are high and stable.[10]
2. A standard finding in iterated social dilemma experiments is that without communication cooperation gradually declines.[11] Cooperation rates peak after communication, *even when it takes place after a few rounds of declining cooperation*.[12]
3. 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, for example, designed a prisoner's dilemma game with two

periods of 10 stages each. The experiment included three conditions: (1) no communication, (2) communication before the second period, but not before or during the first period, and (3) communication before the first period, but not before the second period. They found that, in the third condition, cooperation in the first period was almost 100 percent and carried over to the second period (85 percent on average). In the first condition (no communication), cooperation started at 50 percent, but quickly deteriorated to 10 percent.[13]

4. When there are two separate groups, the *carryover effect* extends to the out-group.[14] Orbell, Van de Kragt, and Dawes 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 percent of the subjects gave to the out-group after discussion significantly more than the 30 percent contribution rate in the control, no-communication condition (but still less than the 79 percent contribution rate obtained when communication was allowed and the contributions went to the original in-group beneficiaries). In a similar, step-level, public-good experiment, Van de Kragt et al. show that even when the gains from cooperation are not enjoyed by the in-group, communication increased contribution rates to a sizable 76 percent, whereas the cooperation rate when communication was not allowed was a meager 30 percent.[15]
5. The communication effect lasts even when agents are made aware of incentives to defect. As Kerr et al. found in a step-level, public-good experiment, subjects continued to cooperate after discussion even after being focused on the short-term monetary losses associated with their choices.[16] The same authors also found that communication promoted cooperation even if agents' choices were not made public and the identities of 'defectors' were kept secret.[17]

After characterizing the relevant features of the communication effect, the next step is to 'decompose' the FtF communication process to identify those features of communication that facilitate cooperation. Dawes identified three elements of FtF communication that make cooperation possible: *identification*, *discussion*, and *commitment*.[18] Experimental results allow us 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. Let us elaborate on these findings.

One possible hypothesis is that the communication effect is due to the *identification* (*humanization*) of other agents that occurs through discussion. Since most of the experiments we discuss involved FtF communication, subjects had a chance to observe each other, form a quick impression about the individuals they were facing, and find (true or imagined) similarities with them. When another subject is perceived as similar to ourselves, even if the similarity is vague and generic, we have a tendency to be kinder and more generous than we would be if the other party were completely anonymous.[19] However, experimental evidence suggests that identification is not sufficient to explain the communication effect. Two experiments show that, indeed, simple identification supports cooperation, but its influence on cooperation is significantly weaker than we observe with interactive communication. In one experiment, Bohnet and Frey found that, in prisoner's dilemma games with varying degrees of social distance, cooperation rates were only 12 percent when subjects were anonymous, 23 percent when subjects identified each other but were unable to communicate, and 78 percent in the FtF communication treatment. Another experiment by Brosig, Ockenfels, and Weimann, which is described later in greater detail, also shows that identification alone induces much less cooperation than other conditions in which two-way communication is allowed.[20]

A second hypothesis about the communication effect attributes it to the cues about the character and motives of agents that are exchanged during *discussion*. Such cues, it is argued, allow agents to form expectations about the future behavior of other subjects. This hypothesis has also been rejected. Experimental data show that *when discussion lacks explicit promising, it can lose its effectiveness in supporting cooperation*. For example, Bouas and Komorita allowed subjects to communicate with each other in a social dilemma game. While in one treatment subjects were allowed to talk about any topic, including the dilemma, in the other treatment they were explicitly forbidden to discuss the dilemma and, by implication, to make promises about future actions. When subjects were allowed to discuss the dilemma and make promises, the cooperation rate was 81 percent. By contrast, when they were allowed to communicate, but not discuss the dilemma, the cooperation rate was only 17 percent.[21]

It has indeed been observed that during FtF communication subjects routinely emphasize the need to make and keep promises to cooperate, and assure other subjects that they will not be taken advantage of.[22] Moreover, subjects often ask one another to make *public* commitments, declare aloud their intentions to cooperate, and even swear to them.[23] The inability to communicate and make promises would thus seem to jeopardize the possibility of cooperation. 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, but even to agents who were not the original recipients of the pledge and did not promise anything in return.

2. The communication effect and the focus theory of norms

The FtF communication effect has recently been explained by Bicchieri in terms of her theory of social norms. This explanation is consistent with and supported by the experimental findings described above. According to Bicchieri, communication about the dilemma can *focus* agents on pro-social norms (particularly the norm of promise keeping); when it does so, communication generates a normative environment that is conducive to cooperation. Before further exploring the communication effect in computer-mediated environments, it is necessary to sketch the conditions for the existence and activation of social norms, upon which the communication effect is arguably grounded. A fuller exploration of these topics is beyond the scope of this article.[24]

By ‘social norms’ we refer to informal behavioral rules that are not supported by formal sanctions. Moreover, we take social norms to be defined by the dispositions that people have toward them. That is, what is a social norm for one person may be a personal norm for another, and just a convention for a third. 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 on expecting most other people to keep their promises as well, and on the belief that one is expected to fulfill one’s promises, too. When we talk of a *social* norm of promise keeping, we always refer to a behavioral rule that is followed only if certain conditions are met, but may otherwise be disregarded. As we shall see in the following discussion of social dilemmas, this view of social norms is crucial for explaining the effects of communication on cooperation. Cooperation does 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.

What follows is a brief definition of the conditions under which a social norm exists and is followed:[25]

> Let R be a *behavioral rule* for situations of type S, where S can be represented as a mixed-motive game. We say that R is a social norm in a population P if there exists a sufficiently large subset $P_{cf} \subseteq P$ such that, for each individual $i \in P_{cf}$:
>
> *Contingency*: i knows that a rule R exists and applies to situations of type S, and
>
> *Conditional preference*: i prefers to conform to R in situations of type S on the condition that:
>
> 1. *Empirical expectations*: i believes that a sufficiently large subset of P conforms to R in situations of type S and either
> 2. *Normative expectations*: i believes that a sufficiently large subset of P expects i to conform to R in situations of type S or
>
> 2′. *Normative expectations with sanctions*: i believes that a sufficiently large subset of P expects i to conform to R in situations of type S, prefers i to conform, and may sanction behavior.

A social norm R is *followed* by population P if there exists a sufficiently large subset $P_f \subseteq P_{cf}$ such that, for each individual $i \in P_f$, conditions 1 and either 2 or 2′ are met for i and, as a result, i prefers to conform to R in situations of type S.

The existence and motivational force of a social norm thus depend upon a sufficiently large number of people having the right kind of empirical and normative expectations, as well as conditional preferences for following the norm, given those expectations. It seems clear that, though a norm may exist, it may not be followed in all circumstances to which it can apply. There are several reasons why this may happen. On the one hand, an agent may not be focused on specific situational cues that make a particular norm salient. Even if a norm has become salient, an agent may have reasons not to expect other agents to follow it. Finally, an agent may not believe that others expect her to follow the norm, or will punish her transgression.

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.[26] Another is to be exposed to written or verbal content that 'calls to mind' a specific norm.[27] In laboratory experiments, norms can be primed by the experimenter's instructions.[28] Direct communication, however, is one of the most effective mechanisms for focusing people on social norms. Indeed, FtF communication *about the dilemma* seems to have a strong focusing effect. We have already stressed that communication is effective when the topic discussed is the social dilemma, and that discussion results in pledges to cooperate. It might be argued that promises exchanged in a one-shot context, where it is known that subsequent decisions will be made anonymously, are just cheap talk and should not be expected to be kept. Thus, even if FtF communication might succeed in focusing subjects on norms, it remains to be explained why subjects obey them when they could defect with impunity. Bicchieri's theory of norms states that agents have a *conditional preference* for following a norm, provided they have the right kinds of expectations.[29] When promising to cooperate in a dilemma context, what kind of expectations can be generated that would support promise keeping?

It should be noted that exchanging promises is a common, habitual activity, and a frequent experience we share is that people who make promises tend to keep them. In other words, habitual activities are activities that refer to and follow a particular *script*.[30] When we make a promise, we almost automatically prime a script telling us how to behave now and in the future, and what to expect of others involved in the interaction. It has been argued that norms are *embedded* into scripts,[31] and this is the reason why promises in dilemma settings, by priming a familiar script, elicit the kind of expectations that support norm-abiding behavior. Attention to cues is a critical part in the process of recognizing a situation as familiar and priming the appropriate script. If two subjects are presented with the same context, but are focused on different cues, then they will probably interpret the situation very differently. In such case, we could expect different behaviors. It also follows that for two culturally homogeneous subjects to express the same preference (say, for keeping a promise) in a given context, they must pay attention to the same cues and interpret them in the same way. Consequently, the same script will be activated in each of them and elicit the same normative behavior. It is thus extremely important to know as precisely as possible about the conditions under which a given norm becomes operative, that is, when the mapping from a context to a specific interpretation (script) involving a norm will occur.

When agents exchange promises to cooperate in an FtF context, such pledges are supported by contextual cues such as tone of voice, gestures, eye contact, and so on. All these cues contribute to making the situation familiar and the outcome predictable. If agents perceive mutual pledges as true, credible promises, they will feel a normative pressure to keep them, and thus cooperate. Whether we obtain the same result with different types of communication is a question of the greatest importance, given that a great number of interactions now take place over 'impersonal' channels such as the internet.

3. The communication effect in computer-mediated environments

In the first part of the article we argued that the communication effect is rooted in the ability to focus agents on the norm of promise keeping. Consequently, the dilemma situation is perceived as representative of other situations in which agents make and keep their promises. This *scripted* interaction allows agents to form expectations about the behavior of other parties, and beliefs that others expect them to honor their promises as well. All this sets the cognitive groundwork for a cooperative choice. In this part of the article we ask to what extent the communication effect is replicated in social dilemmas in which communication is computer mediated. Do promises have the same credibility and effects when delivered through computers and not face to face?

In typical noncomputer-mediated interactions, the costs of solving social dilemmas by mutual promises are quite high, as coordinating mutual promises among group members may be difficult, especially as group size grows large and agents are geographically dispersed. The costs of organizing a 'social contract' dramatically decrease online.[32] At the same time, experimental work demonstrates that using computer-mediated communication instead of FtF communication can hamper the generation of normative settings in which promises are perceived as reliable. There are relatively few social dilemma experiments that study the effects of communication in computer-mediated contexts.[33] Let us consider their findings in some detail, in order to pinpoint the crucial differences between computer-mediated and FtF communication effects on normative behavior.

A notable experiment by Brosig, Ockenfels, and Weimann nicely captures some of the points we make below, and is worth reviewing in some detail.[34] The experiment is about a repeated public-good game with seven different conditions, most of which involved pre-play communication. The only difference between the communication conditions was the *medium* of pre-play communication. The seven conditions included (1) a control, no-communication condition and (2) an identification condition in which players saw one another for 10 seconds, but could not talk. Two of the other conditions involved unidirectional communication, in which players were exposed to communication from others, but could not actively participate. These were, respectively, (3) a 'lecture' condition and (4) a 'talk-show' condition.[35] In another condition, (5) the group communicated through an audio-conference system. In the last two conditions, groups communicated using either (6) a videoconference system or (7) a 'table conference' (subjects sat around the same table and could talk to each other for up to 10 minutes).

The authors found that different media produced distinct communication effects, and that contributions in the videoconferencing and table-conference conditions were significantly higher than in other conditions. The seven communication conditions were divided into three subgroups according to the extent and stability of cooperation. Each subgroup includes similar conditions, but such conditions significantly differ from those of other subgroups. The three subgroups are as follows:

1. Control and identification conditions were characterized by relatively low initial-contribution rates (less than 50 percent), but relatively stable cooperation (decline of less than 20 percent).
2. The audio-conference condition, as well as the two conditions of unidirectional communication ('lecture' and 'talk show'), were characterized by intermediate initial-contribution rates (between 50 and 60 percent), but relatively unstable cooperation (decline of more than 30 percent).
3. The videoconference and 'table-conference' conditions were characterized by high initial-contribution rates (more than 90 percent) and relatively stable cooperation (decline of less than 20 percent).

It was found that the communication effect was present in *all* communication conditions, but different computer-mediated communication (CMC) channels produced different levels of contributions to the production of the public good. In the absence of continuous communication there was a decline in cooperation rates in all conditions, although the extent of the decline varied across conditions.

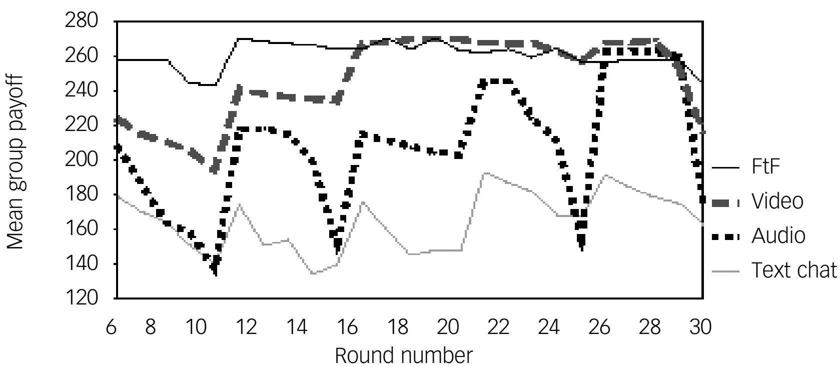
Other, related experiments show a similar pattern. Frohlich and Oppenheimer ran a repeated public-good game with three conditions: (1) a no-communication control condition, (2) an email communication condition, and (3) an FtF communication condition.[36] The experiment consisted of two stages. In the first eight rounds of play (stage 1), communication was allowed before every round. In the remaining seven rounds (stage 2), communication was not permitted. At the end of each round, payoffs were calculated and displayed. Subjects were unaware of how many rounds would be played in either stage, and did not know at the outset that there would be two stages of play. Frohlich and Oppenheimer found significant differences in contribution rates between the three conditions. In the first stage of the game, contribution rates in the no-communication control condition were 29 percent, significantly less than in the email communication condition (76 percent) and the FtF communication condition (99 percent). The communication effect was present in both communication conditions, although email communication generated significantly less cooperation than FtF communication.

In the second stage of the experiment, when communication was not allowed, contributions in all the previous communication conditions declined dramatically. While there were significant differences between the cooperation rates of the groups that communicated and the group that did not communicate in stage 1, there were no statistically significant differences between the two communication conditions in stage 2, and by round 10 the cooperation rates of no-communication and email communication were nearly identical.

Bochet, Page, and Putterman got similar results in a 10-round, public-good game with four conditions: (1) a no-communication control; (2) a 'numerical cheap talk' (NCT) condition, in which subjects could anonymously send via computer a nonbinding message about how much money they intended to contribute before each period; (3) a chat-room condition, in which subjects could participate in unrestricted communication every few periods; and (4) an unrestricted pre-play FtF communication condition (with no subsequent communication during the game).[37] After subjects made their contribution decisions, they learned about others' decisions and their own earnings. The authors found no significant difference in cooperation rates between the NCT condition and the no-communication control. This is not surprising since, in the 'thin' communication condition of NCT, there was a lack of correspondence between proposed and actual contributions, and this brought about a gradual decline in cooperation rates. Contributors in the chat-room condition were significantly lower than in the face-to-face condition, but still significantly higher than in the no-communication baseline and the NCT conditions.

To tease apart the effects of promising from those of interactive communication, in a follow-up paper Bochet and Putterman added another condition (similar to NCT) in which anonymous subjects, before making a contribution, could make a nonbinding *promise* after the initial round of cheap talk (consisting in the exchange of simple numerical statements about future monetary contributions).[38] Before making a choice, players were informed about the pledges of all group members. There were no significant differences in actual contributions between the 'numerical cheap talk with promising' (NCTwP) condition and the no-communication condition. This result suggests that the normative effect of promising depends upon the context in which promises are made; small groups (in chat rooms) or face-to-face communication allow for interactive communication, thus providing subjects with opportunities to check each other's intentions, threaten retaliation, and coordinate action that are unavailable in 'thin' communication conditions.

Further support for the latter hypothesis comes from an experiment by Bos et al.[39] In this experiment, subjects played 30 rounds of a social dilemma game, and communication was allowed every five rounds using, respectively, text chat, audio, video, or FtF. It was found that groups using videoconferencing to communicate achieved nearly the same levels of contributions as the FtF groups, although it took them more time to coordinate on cooperation. There were significant differences in cooperation rates between the various conditions (see Figure 1). As the graph shows, whereas the FtF and videoconferencing groups achieved 'smooth' levels of cooperation, there were sharp increases in cooperation in the text and audio groups after discussion, followed by equally sharp declines. Since the outcome of each round was public knowledge, the authors attribute these fluctuations to defections followed by retaliation in the between-communication rounds and, conversely, increases in cooperation rates immediately following communication. Defections were minor in the FtF and videoconferencing conditions.



Source: Bos et al., 'Being There versus Seeing There: Trust via Video'.

Figure 1 **Contribution patterns in a game with FtF, video, audio, and text chat conditions**

As in many other repeated social dilemmas, the decline in cooperation rates can be attributed both to a better understanding of the agents' incentives and to a weakening in the motivational force of social norms. These two explanations are not mutually exclusive: understanding the strategic properties of the game is triggered by lower than expected cooperation rates, which in turn make agents reconsider social norms. For example, Andreoni claims that many agents learn the dominant strategy before they implement it, and do so only after trying to engender cooperation. The decline in cooperation rates represents, then, what Andreoni calls 'frustrated attempts at kindness'.[40] We believe that a large number of agents start the game with a 'default' cooperative norm. Such a norm is supported by empirical and normative expectations.[41] However, if such expectations are violated, we predict that norm-following agents will revert to more selfish behavior. This is precisely what happens in repeated social dilemmas in the absence of communication or in the presence of 'thin', and thus ineffective, communication channels.

The fragility of text-based cooperation noted by Bos et al. is evident in another experiment by Rocco on common-pool resource depletion.[42] The experiment had two communication conditions, an FtF condition and a 'mailing list' condition, in which everyone could access everyone else's messages (as in a newsgroup). In both conditions, communication was only allowed after the 10th, 15th, and 20th rounds of the game. The author reports that the FtF groups quickly converged on cooperation, which remained stable with no defections occurring. By contrast, in the 'mailing list' condition subjects failed to reach effective cooperative agreements. Although 83 percent of the electronic communication periods resulted in some of the parties agreeing to cooperate, only 27.8 percent of them involved unanimous agreement.[43] It might be argued that the lack of unanimity caused most agreements to be immediately disregarded and followed by massive defections, in that the expectation of compliance with the terms of the agreement could not be sustained. Rocco, however, noted that in the 'mailing list' condition *every* agreement was disregarded *irrespective* of the presence or absence of unanimity. Rocco attributes this dismal result to the breakdown of traditional conversation rules and their replacement by a 'chaotic' flow of communication, as well as to the difficulty that electronic communication presents for the emergence of leaders who organize and focus discussion, as often happens in experiments with FtF communication.[44]

To check whether the small number of communication periods was at least in part responsible for the dismal results of Rocco's 'mailing list' conditions, we should look at similar experiments in which subjects had a greater number of communication periods available to them. For example, Jensen et al. allowed subjects to communicate continuously in an iterated prisoner's dilemma game, using audio (a speakerphone), text-to-speech (TTS) interface, and simple text messages. There was also a control condition without communication. Cooperation rates were higher than in Rocco's experiment, though the medium of communication made a big difference. Indeed, there were statistically significant differences in cooperation rates between the audio condition and both the text condition and the no-communication condition.[45]

In another experiment by Zheng et al., agents played a variant of the prisoner's dilemma game. The experiment included five conditions in which subjects were involved in different pre-play 'social activities'. During the game itself, subjects were allowed to communicate every five rounds using an internet-based chat program. The communication effect was present in all conditions, and was especially evident after the first round of communication. 'Chatting' supported an increase in contributions, even after contributions declined in the rounds with no communication. The authors report significantly higher cooperation rates (with respect to the no-communication condition) when subjects communicated FtF before the game or participated in internet-based 'social chat'. However, if in the pre-play period subjects only saw photos or written personal descriptions of their partners, cooperation rates were not significantly higher than in the control, no-communication condition.[46]

This survey of experimental results enables us to draw some initial conclusions about the communication effect in computer-mediated environments. Here is a brief summary of the main findings about the presence and magnitude of the CMC effect in social dilemma experiments:

1. The communication effect is still present in computer-mediated environments: CMC produces higher cooperation rates than equivalent environments in which communication is not allowed.
2. The communication effect varies in degree according to the richness of the communication channel. Videoconferencing produces cooperation rates very close to FtF communication, whereas text-based communication produces substantially 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*.
3. When using CMC to communicate in social dilemmas, communication is normatively charged, and even more forceful than FtF communication. This could be explained by the need to 'compensate' for the lack of contextual cues in computer-mediated environments.[47]
4. Compared to FtF communication, it takes more time to establish cooperation, especially when using 'poorer' CMC channels. As in FtF communication, cooperation deteriorates over time in the absence of continuing communication. After communication resumes, cooperation rates improve again (which also occurs in FtF communication).
5. 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.

4. Communication of mutual promises

In this section, we will try to account for the differences between computer-mediated and FtF communication effects, and conclude that some CMC environments cannot effectively focus agents on the relevant social norms, thus preventing them from developing expectations about the future actions of others as well as about others' normative expectations about their own actions. Since in face-to-face communication the cooperative result is obtained through the exchange of promises, we should pay special attention to the conditions under which promises can be made (and be credible) in different communication conditions. Unfortunately, much of the literature on promising has failed to focus on the environmental conditions under which promising occurs, but, as we shall see below, such conditions are crucial in conveying the credibility of mutual intentions. A similar consideration applies to social norms in general. Different contexts hamper or promote focusing on the relevant norms and forming those expectations that are so crucial in supporting norm-abiding behavior.[48]

It is almost trivial to say that social interaction is grounded on a massive amount of collaborative work and common knowledge that are ordinarily taken for granted.[49] Communication is a collective activity aiming, among other things, at establishing a sufficient degree of 'common grounding' among communicators, that is, a common belief that all communicators understand each other and share common premises and meanings. Clark and Brennan argue that the collaborative effort to 'ground' communication includes, among other things, variables such as *co-presence* (interlocutors share the same physical space), *visibility* (communicators are visible to one another), and *audibility* (communication can be performed by verbal utterances). They also mention important conversational factors, such as *contemporaneousness* (messages are received roughly at the same time they are produced), *simultaneity* (both communicators can send and receive messages), *sequentiality* (there exist known 'protocols' that communicators use to take turns while communicating), as well as the possibility of *reviewing and revising* messages before sending them.[50] Not all of these variables are present when communication occurs, but most of them are present in FtF communication, whereas many are absent in 'thinner' computer-mediated communication channels. As we shall see below, some of these factors are more important than others when agents communicate with each other about the dilemma and promise to cooperate.

We know that when individuals face new situations, they tend to categorize them immediately as relevantly similar to other, more familiar situations, and then apply rudimentary knowledge structures (scripts or schemata) to the new situation in order to understand, explain, and predict what has occurred or is going to occur as well as to fine-tune their behavior.[51] Individuals are obviously familiar with the structure and communication of promises delivered FtF, and tend to expect that other agents will fulfill their promises even in CMC dilemma settings. That is, when a norm of promise keeping is activated, mutual expectations, beliefs, and a conditional preference for following the norm will also be simultaneously activated in the vast majority of participants.[52] However, when the environment and the means of communicating promises differ significantly from familiar settings in which promises are usually made, agents may become focused on the 'poverty' of the normative environment, fail to develop expectations about the future actions of promise-makers, and as a result may decline to cooperate themselves. FtF promises are often preceded by long discussions and are supported by a range of nonverbal cues. Videoconferencing roughly approximates the reliability of promises delivered when individuals are in the same place. Removal of visual or auditory channels of communication results in environments that are further removed from those in which agents regularly make promises and expect promises to be kept.[53]

To address the differences between online and face-to-face promising and their influence on behavior, let us further decompose the mechanism of promise-making. In particular, we shall concentrate on three aspects of 'commitment production': (1) the capability to coordinate mutual promises, (2) the credibility of the act of mutual promising, and (3) the possibility of attaining public knowledge about the mutuality of promises. We argue that these three aspects of 'commitment production' become problematic in computer-mediated environments. As a consequence, people will be more inclined to question the credibility of online promises, with detrimental consequences for cooperation.

4.1 How to coordinate a 'social contract'

We know that, in the absence of institutions that create and enforce rules, it can be difficult to engage in mutual promising (let alone to converge on a consensual normative understanding of the dilemma). Social norms are endogenously evolved informal institutions that offer a behavioral solution (for example, 'keep your promise') when certain conditions are met (see Section 2). In particular, individuals must possess the right kind of normative and empirical expectations in order to prefer conformity to the norm. In a situation in which many options are possible, focusing on mutual promising and following through requires the presence of certain contextual conditions that may be absent in computer-mediated communication. In particular, we want to consider, first, the *capability to coordinate mutual promising*, something that is often overlooked in favor of the more pressing problem of credibility.

In CMC conditions, two main complications are likely to occur. The first hurdle is that familiar conversational structures are altered (especially in asynchronous communication). Turn taking, sequencing, and alternation become problematic. Overlapping utterances and communication delays frequently occur. Furthermore, often no immediate feedback is available to the parties involved.[54] Therefore, when communication is computer mediated, coordinating a solution to a complex task, such as agreeing on the content of promises and then engaging in mutual promising, can be exceedingly difficult.[55]

The second obstacle is that the absence of FtF communication may prevent the emergence of leadership. Studies of communication in social dilemma experiments have reported that some form of 'leadership' is instrumental for coordinating mutual promising and establishing a 'common script' among participants. For example, Orbell, Van de Kragt, and Dawes maintain that 'in some groups some individuals pushed the idea of getting explicit promises from everybody – often with success. When no individual was willing or assertive enough to do this (or considered the possibility), promise making seemed much less likely to happen.'[56] Self-selected 'leaders' dominated the discussion and advocated specific strategies. In their absence, agents found it difficult to reach agreements and often ended the discussion ahead of time.

The breaking of familiar conversation rules and the difficulty of creating leadership make it harder to coordinate computer-mediated promises. Unfortunately, experiments that directly test the consequences of these difficulties for achieving cooperation are still to come.

4.2 The credibility of promises

The second aspect of 'commitment production' that is called into question in CMC is the *credibility* of promises. Though a traditional rational choice argument makes *any* promise to cooperate in a one-shot, social dilemma game non-credible, we have plenty of evidence of promise-making and promise keeping in such games. Bicchieri explains such behavior as normatively oriented. Mutual promising involves an exchange of pledges (to cooperate, in our examples), and the norm of promise keeping on which we thus focus is supported by the expectation that our party will keep it, as well as by the belief that our party expects us to keep our promise. Not everyone is equally sensitive to a norm, but experimental data show quite conclusively that for most people a norm of promise keeping is meaningful and will be adhered to, provided the relevant reciprocal expectations are fulfilled.[57] When promises are made in FtF contexts, a variety of cues that are available allow subjects to assess intentions and form expectations about each other, all of which can make their mutual promises credible. Such indicators include *visual cues* (that is, 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 frequently correlated by agents with trustworthiness, and their presence or absence can have important motivational consequences via the formation (or impairment) of mutual expectations of promise-keeping behavior.

When communication is computer mediated, such cues are often restricted or entirely absent (and especially so in 'thinner' communication channels); consequently, a range of indicators about intentions is restricted or unavailable. Indeed, experimental results show that when promises are not supported by the subtle cues and richer language structure that typically back FtF promising, they are perceived as less credible, less binding, and thus result in higher defection rates.[58] We may conclude that when intentions become opaque, cooperation falters. It must be noted that in computer-mediated communication it is not just the case that *less* people engage in promising. It is also the case that many of those who promise *do not keep* their promises. Since agents are not induced in any way to promise anything, promising in this case may represent strategic behavior (to send a cooperative signal with the intention to defect). However, if indeed agents sometimes engage in strategic promising, it must be that some of the characteristics of FtF contexts discourage strategic promising. Indeed, FtF communication *induces* agents to promise more and to act consistently with their pledges.

The experimental literature is rife with cases in which subjects make promises, but do not keep them. It is worth noting that in none of these cases is there face-to-face communication. Palfrey and Rosenthal, for example, ran a step-level, public-good experiment that included one round of pre-play communication in which subjects could announce, using their computer terminals, their intentions to contribute or not, knowing that their announcement was nonbinding. They found no significant difference in contribution rates between this condition and a no-communication control condition. Bochet, Page, and Putterman similarly found no significant difference in cooperation rates between a no-communication control condition and a 'cheap talk' condition in which subjects could send via computer a short numerical message before each period of play. Likewise, Chen and Komorita show that when subjects wrote their pledges on a piece of paper, contributions did not significantly differ from the no-pledge control condition. Finally, consider two experiments quoted by Ostrom that faithfully replicated the designs of previous social dilemma experiments with a single modification: discussion and promises did not take place FtF, but were computer mediated. In both experiments, cooperation rates were substantially lower when promises to cooperate were generated through computer messages and not FtF.[59] Furthermore, a large discrepancy between promises and subsequent behavior was observed, in that the majority of pledges to cooperate resulted in contributions much smaller than was promised.

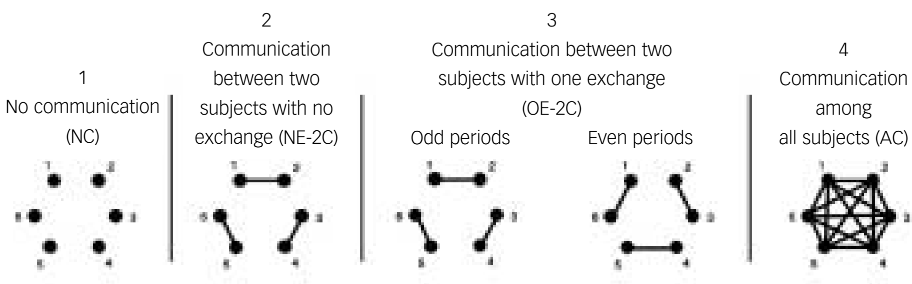
In all these cases, it is realistic to assume that agents did not perceive others' promises as credible, and may have even been aware that their own promises might have been regarded with skepticism by the other parties. In this case, both empirical and normative expectations falter, and it is not surprising that one should not feel bound by a norm of promise keeping that almost no one expects or is expected to follow.

4.3 Making promises public knowledge

A third, important part of the mechanism of 'commitment production' is the presence of an *efficient information-dissemination system* that quickly, and with a minimum amount of noise, spreads information about the number of other agents promising to contribute. In social dilemma games, the final outcome depends on the number of agents who choose to cooperate. This number may vary with the specifics of the game, ranging from the case in which just one defection dooms the group to cases in which there is a minimum sufficient number of cooperators. In the latter case, some defections can be tolerated, but it is of course impossible to tell whether one's defection will be critical. In all cases, it is crucial to expect a sufficient number of players to choose cooperatively, otherwise one's cooperative choice might end in a net loss. *Public promising* is one way for conditional cooperators to generate and develop expectations that others will cooperate as well. Even if promising is not unanimous, there usually is a perceived consensus among agents about the appropriateness and expected levels of cooperation.[60] Such perceived consensus is difficult to create in the absence of efficient dissemination of information about the extent of promising. Computer-mediated communication usually lacks such public knowledge features; the potential for noisy communication is much higher in CMC contexts, and it is difficult to assess how many players chose to cooperate. When information about the extent of promising does not reach group members, there is much uncertainty about the intentions of other parties and, as we saw in Section 3, the conditions for norm compliance are missing.

One interesting question to explore is whether *partial communication* in FtF contexts would be sufficient to generate cooperative results. Though the typical experimental FtF communication involves a small group of subjects, in many real-life situations we encounter, communication takes place in large-group settings. In such circumstances, public information about promising may or may not be possible. What is feasible, however, is the formation of small subgroups in which subjects can easily and efficiently communicate with each other, and make mutual promises. As Figure 2 shows, however, there are many possible ways in which a small group of individuals may connect: 'no communication' to 'all communicate with all', with many intermediate degrees of connection among group members. The intermediate cases are the most common ones, and the most interesting to study. When every group member does not communicate with every other member, but just with a subset, is there a critical number of 'links' below which the communication effect disappears?

There is some experimental evidence showing that what matters to efficient communication is not just the size of the group, but rather the number of connections among agents. For example, we know that, in FtF encounters, the



Source: Kinukawa et al., 'Partial Communication in a Voluntary-Contribution-Mechanism Experiment'.

Figure 2 **Communication treatments**

communication effect persists even when communication is partial (some agents could not communicate with others), as well as when agents communicate in subgroups. Braver and Wilson found that when players discussed the dilemma in *subgroups*, cooperation rates reached 75 percent, significantly more than the 48 percent cooperation rate obtained in the absence of communication. More recently, Kinukawa, Saijo, and Une established that communication effectiveness is positively correlated with the *connectivity* of the communication network of agents involved in a public-goods dilemma (see Figure 2).[61] Their experiment included four treatments, each consisting of 10 repetitions. In Treatment 1, no communication was allowed, and contributions decreased over time from 29 percent to 20 percent. In Treatment 2, FtF communication was allowed in dyads, and contributions decreased from 34 percent to 19 percent. There was no statistically significant difference in contribution rates between this treatment and the no-communication treatment. In Treatment 3, each agent was allowed to alternate communicating FtF with only one of two other agents in such a way that the network of communication among all subjects indirectly connected over time.[62] Interestingly, in the third treatment contribution rates increased over time from 33 percent to 58 percent. In a fourth treatment, in which full FtF communication was allowed, contributions were almost 100 percent from the outset.

These results suggest that even when communication is fragmented, it can still have beneficial effects on cooperation. Depending on the possibilities of communication across fragments, as well as the specifics of the social dilemma that is analyzed, it seems possible to generate the three crucial aspects of 'commitment production': *coordination* of promising, *credibility* of promises, and *public knowledge* of the pledges that have been made. Further studies are needed to explore if the same positive outcomes on cooperation occur when such fragmentary discussions are computer mediated.

5. Conclusions

We argued that communication affects cooperation in social dilemmas because it focuses agents on pro-social norms, notably the norm of promise keeping, and generates empirical and normative expectations that support a conditional choice to keep one's promise. We also argued that the communication effect can be present even when a computer interface replaces face-to-face communication. Yet not every type of computer-mediated communication lends credibility to promises. Richer media (especially videoconferencing systems) allow agents to attain public knowledge of their respective promises, as well as to perceive promises as credible and thus feel it is safe to cooperate. 'Thinner' communication channels, instead, may fail to convey enough contextual cues to generate the right kind of expectations that support norm-abiding behavior. Computer-mediated promises may be perceived as less credible than their face-to-face counterpart, and the possibility of producing a collective commitment and making it public knowledge are further impaired by the difficulty of establishing leadership, organizing group discussion, and disseminating information about its content.

The study of computer-mediated communication and of the conditions under which mutual promising and collective commitments are likely to take place is the subject of ongoing research. Computer-mediated communication (and particularly internet communication) allows inexpensive, large-scale discussions that could ultimately support mutually beneficial cooperative solutions in social dilemmas. Though the difficulties of generating and communicating credible promises online can impair the effectiveness of the communication effect, it is worth exploring if and how enriching media channels, or breaking large groups into frequently interacting small subgroups, may have a positive effect on agents' willingness to cooperate.

Appendix

Table A1 **Specifications of the relevant communication conditions in the surveyed experiments**

Studies	Communication conditions
Bochet and Putterman[63]	(1) An NCT condition in which subjects could send a nonbinding numerical declaration before each round; (2) same as NCT, but subjects added a nonbinding statement: subjects were asked to choose between the statement 'I do not wish to make a promise at this time' and the statement 'I promise to contribute to the group account this period', in which case they selected a number between zero and 10.
Bochet et al.[64]	(1) An NCT condition in which subjects could send a nonbinding numerical declaration before each round; (2) a chat-room condition in which subjects could participate in unrestricted communication every few rounds (before rounds 1, 4, 7, and 10 of 10); (3) a 5 min., pre-play, URFFC (no subsequent communication).
Bohnet and Frey[65]	10 min., pre-play URFFC.
Bornstein[66]	5 min., pre-play URFFC.
Bos et al.[67]	Communication was allowed every five rounds; communication conditions include: text chat, audio, video, and FtF. (*)
Bouas and Komorita[68]	10 min. of either (1) pre-play URFFC or (2) pre-play FtF 'irrelevant communication' (discussion of the game was forbidden).
Braver[69]	10 min., pre-play URFFC.
Braver and Wilson[70]	10 min., pre-play URFFC.
Brosig et al.[71]	10 min., pre-play unrestricted communication using (1) an audio-conference system or (2) a videoconferencing system or (3) FtF.

continued

Table A1 **cont.**

Studies	Communication conditions
Cason and Khan[72]	4 min. URFFC during the game; rounds in which communication was permitted varied across sessions.[73]
Chen[74]	Subjects used computer interfaces to declare intended contributions before each round. (*)
Chen and Komorita[75]	Subjects declared their intended contributions on a piece of paper prior to each round. (*)
Dawes et al.[76]	10 min. of either (1) pre-play URFFC or (2) pre-play FtF 'irrelevant communication' (discussion of the game was forbidden).
Frohlich and Oppenheimer[77]	(1) An email communication condition and (2) a URFFC condition. Communication was allowed before each round, with the exception of the first eight rounds. (*)
Hackett et al.[78]	10 min. URFFC after the 10th round, then 3 min. before each round.
Isaac and Walker[79]	4 min. URFFC was allowed either before the 1st round or before the 11th round (see Figure 1).
Jensen et al.[80]	Unrestricted communication before each round, using audio, TTS interface, and text messages. (*)
Jerdee and Rosen[81]	1 min. URFFC after each round.
Kerr and Kaufman-Gilliland[82]	5 min., pre-play URFFC.
Kerr et al.[83]	5 min., pre-play URFFC.
Kiesler et al.[84]	Subjects had to declare a binary choice (to contribute or not) in response to identical and pre-generated sequences of communication and choices, conveyed either FtF by a human or by a computer using text, audio, or 'moving face' interfaces.
Kinukawa et al.[85]	Three URFFC conditions of 2.5 min. before each decision round: (1) communication in predetermined dyads, (2) communication in dyads with one of two partners, when the partner changed after each period, and (3) group communication.
Kurzban et al.[86]	Group members used a real-time, computerized, voluntary-contribution mechanism to declare their intended contribution. All subjects could see the current declarations of other subjects, but could not identify other subjects. Subjects could use a 90 second interval to make or alter declarations. Their actual contribution in each round was their declaration when the countdown clock reached zero.
Orbell et al.[87]	10 min., pre-play URFFC.
Ostrom and Walker[88]	4 min. URFFC prior to each round, after the 10th round.
Ostrom et al.[89]	(1) 'One-shot communication': 10 min. URFFC at the end of the 10th round. (2) 'Repeated communication': URFFC each round after the 10th round. (*)

continued

Table A1 **cont.**

Studies	Communication conditions
Palfrey and Rosenthal[90]	Subjects sent a pre-play signal: I (intend to spend) or NI (intend not to spend).
Rocco; Rocco and Warglien[91]	(1) 10 min. URFFC after rounds 10, 15, and 20 and (2) 30 min. unrestricted communication through a mailing list, after the same rounds.
Schmitt et al.[92]	10 min. URFFC after the 10th round, then 2 min. before each round.
Van de Kragt et al.[93]	10 min., pre-play URFFC.
Van de Kragt et al.[94]	10 min., pre-play URFFC.
Wilson and Sell[95]	Subjects could use a computerized interface to announce their intended contributions before each round.
Zheng et al.[96]	Five conditions in which before play (1) no communication was allowed, (2) personal information about partners was presented, (3) photo of partner was presented, (4) 15 min. unrestricted text chat was allowed, and (5) 10 min. URFFC was allowed. During the game, text-chat communication was allowed every five rounds in all conditions.

Notes: Face to face (FtF), numerical cheap talk (NCT), text to speech (TTS), unrestricted face-to-face communication (URFFC), (*) the length of the communication period is not specified. Superscript figures refer to endnotes.

notes

1. Cristina Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms* (New York: Cambridge University Press, 2005); Cristina Bicchieri, 'Covenants Without Swords: Group Identity, Norms and Communication in Social Dilemmas', *Rationality and Society* 14 (2002): 192–228.
2. A typical social dilemma experiment uses the mixed-motive structure of an *n*-person prisoner's dilemma to study choice behavior. It is individually best for each subject to keep her money (that is, not to contribute to public-good production), but all are better off if everyone makes a cooperative decision.
3. John O. Ledyard, 'Public Goods: A Survey of Experimental Research', in *Handbook of Experimental Economics*, edited by John H. Kagel and Alvin E. Roth (Princeton, NJ: Princeton University Press, 1995), pp. 111–94; David Sally, 'Conversation and Cooperation in Social Dilemmas', *Rationality and Society* 7 (1995): 58–92. For additional papers documenting and analyzing the communication effect, see Bicchieri, 'Covenants Without Swords: Group Identity, Norms and Communication in Social Dilemmas'; Shirli J. Kopelman, Mark Weber and David M. Messick, 'Factors Influencing Cooperation in Commons Dilemmas: A Review of Experimental Psychological Research', in *The Drama of the Commons*, edited by Elinor Ostrom, Thomas Dietz, Nives Dolsak, Paul C. Stern, Susan Stonich and Elke U. Weber (Washington, DC: National Academy Press, 2002), pp. 113–56; Anisha Shankar and Charles Pavitt, 'Resource and Public Goods

Dilemmas: A New Issue for Communication Research', *Review of Communication* 2 (2002): 251–72.

4. The experimental laboratory is an indispensable tool to study the communication effect. It enables precise specification of the choice setting, allows manipulation and control of a variety of intervening variables, and rules out competing hypotheses by conducting multiple treatments. However, laboratory experiments involving communication face an external validity problem, in that the conditions for communication are provided by the experimenter and (in most cases) are costless for the subjects. By contrast, in 'real-world' social dilemmas the costs of communication (and especially discussion) can be quite high. Providing a public space for communication is a public good in itself. Messick and Brewer state that outside the laboratory, 'the implications of [the communication effect] are limited . . . in that for many real-world dilemmas, such direct communication among group members is not an available solution to the problem. Most social dilemmas involve large collectives that are extended in time and space, offering little or no opportunity for group members to communicate or negotiate a solution to the choice problem.' See David M. Messick and Marilyn B. Brewer, 'Solving Social Dilemmas', in *Review of Personality and Social Psychology*, Vol. 4, edited by Ladd Wheeler and Phillip Shaver (Beverly Hills, CA: Sage, 1983), p. 23. Internet communication facilitates inexpensive and widely accessible public communication channels in many dilemma situations, and therefore may potentially solve the provision problem and with it the external validity problem. See Azi Lev-On, 'Web-Based Collaboration and the Organization of Democracy', PhD dissertation (New York: New York University, 2005).
5. Robyn M. Dawes, Jeanne McTavish and Harriet Shaklee, 'Behavior, Communication, and Assumptions about Other People's Behavior in a Commons Dilemma Situation', *Journal of Personality and Social Psychology* 35 (1977): 1–11. In a commons dilemma, groups share a scarce resource which individual members can harvest. It is individually beneficial to appropriate a large amount of the common resource, but all are better off if everyone takes a small portion of the resource, so as not to deplete it.
6. Elinor Ostrom and James M. Walker, 'Communication in a Commons: Cooperation Without External Enforcement', in *Laboratory Research in Political Economy*, edited by Thomas R. Palfrey (Ann Arbor: University of Michigan Press, 1991), pp. 287–322. For similar results, see Elinor Ostrom, James Walker and Roy Gardner, 'Covenants With and Without a Sword: Self-Governance is Possible', *American Political Science Review* 86 (1992): 404–17; Roy Gardner, Elinor Ostrom and James Walker, 'Social Capital and Cooperation: Communication, Bounded Rationality, and Behavioral Heuristics', in *Social Dilemmas and Cooperation*, edited by Ulrich Schulz, Wulf Albers and Ulrich Mueller (Heidelberg: Springer, 1994), pp. 375–412.
7. Norman Frohlich and Joe Oppenheimer, 'Some Consequences of E-mail vs. Face-to-Face Communication in Experiment', *Journal of Economic Behavior and Organization* 35 (1998): 389–403.
8. Gary Bornstein, 'The Free Rider Problem in Intergroup Conflicts over Step-Level and Continuous Public Goods', *Journal of Personality and Social Psychology* 62 (1992): 597–606; Steven Hackett, Edella Schlager and James Walker, 'The Role of

Communication in Resolving Commons Dilemmas: Experimental Evidence with Heterogeneous Appropriators', *Journal of Environmental Economics and Management* 27 (1994): 99–126.

9. Based on Bicchieri, 'Covenants Without Swords: Group Identity, Norms and Communication in Social Dilemmas' and Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms*.
10. See Thomas H. Jerdee and Benson Rosen, 'Effects of Opportunity to Communicate and Visibility of Individual Decisions on Behavior in the Common Interest', *Journal of Applied Psychology* 59 (1974): 712–6; Frohlich and Oppenheimer, 'Some Consequences of E-mail vs. Face-to-Face Communication in Experiment'; Ostrom and Walker, 'Communication in a Commons: Cooperation Without External Enforcement'; Pamela Schmitt, Kurtis Swope and James Walker, 'Collective Action with Incomplete Commitment: Experimental Evidence', *Southern Economic Journal* 66 (2000): 829–54; Shinya Kinukawa, Tatsuyoshi Saijo and Masashi Une, 'Partial Communication in a Voluntary-Contribution-Mechanism Experiment', *Pacific Economic Review* 5 (2000): 411–28.
11. For example, Isaac, McCue, and Plott found that by the fourth iteration of a public-good experiment, contribution rates sharply declined from 53 percent to only 16 percent. See R. Mark Isaac, Kenneth McCue and Charles R. Plott, 'Public Goods Provision in an Experimental Environment', *Journal of Public Economics* 26 (1985): 51–74. See also Oliver Kim and Mark Walker, 'The Free Rider Problem: Experimental Evidence', *Public Choice* 43 (1984): 3–24; Robyn M. Dawes and Richard Thaler, 'Anomalies: Cooperation', *Journal of Economic Perspectives* 2 (1988): 187–97; R. Mark Isaac and James M. Walker, 'Communication and Free-Riding Behavior: The Voluntary Contribution Mechanism', *Economic Inquiry* 26 (1988): 585–608; R. Mark Isaac, James M. Walker and Arlington W. Williams, 'Group Size and the Voluntary Provision of Public Goods: Experimental Evidence Utilizing Large Groups', *Journal of Public Economics* 54 (1994): 1–36; Sara Kiesler, Lee Sproull and Keith Waters, 'A Prisoner's Dilemma Experiment on Cooperation with People and Human-Like Computers', *Journal of Personality and Social Psychology* 70 (1996): 47–65; Simon Gächter and Ernst Fehr, 'Collective Action as a Social Exchange', *Journal of Economic Behavior and Organization* 39 (1999): 341–69; Daniel Houser and Robert Kurzban, 'Revisiting Kindness and Confusion in Public Goods Experiments', *American Economic Review* 92 (2002): 1062–9.
12. For example, Ostrom et al., 'Covenants With and Without a Sword: Self-Governance is Possible'; Gardner et al., 'Social Capital and Cooperation: Communication, Bounded Rationality, and Behavioral Heuristics'; Jun Zheng, Elizabeth Veinott, Nathan Bos, Judith S. Olson and Gary M. Olson, 'Trust Without Touch: Jumpstarting Trust with Initial Social Activities', *Proceedings of CHI 2002* (2002): 141–6.
13. Isaac and Walker, 'Communication and Free-Riding Behavior: The Voluntary Contribution Mechanism'.
14. This carryover effect does not seem to work 'in the opposite direction'. When agents hear others talk among themselves about the importance of cooperativeness, or make promises to each other, cooperation rates do not significantly improve. See Robert B. Cialdini, Carl A. Kallgren and Raymond R. Reno, 'A Focus Theory of

Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior', in *Advances in Experimental Social Psychology*, Vol. 24, edited by Mark P. Zanna (San Diego, CA: Academic Press, 1991), pp. 201–34; Norbert L. Kerr and Cynthia M. Kaufman-Gilliland, 'Communication, Commitment and Coordination in Social Dilemmas', *Journal of Personality and Social Psychology* 66 (1994): 513–29; Jeannette Brosig, Axel Ockenfels and Joachim Weimann, 'The Effect of Communication Media on Cooperation', *German Economic Review* 4 (2003): 217–41.

15. J.M. Orbell, A.J.C. van de Kragt and R.M. Dawes, 'Explaining Discussion-Induced Cooperation', *Journal of Personality and Social Psychology* 54 (1988): 811–9; A.J.C. van de Kragt, R.M. Dawes, J.M. Orbell, S.R. Braver and L.A. Wilson II, 'Doing Well and Doing Good as Ways of Resolving Social Dilemmas', in *Experimental Social Dilemmas*, edited by H.A.M. Wilke, D.M. Messick and C.G. Ruttle (Frankfurt am Main: Verlag Peter Lang, 1986), pp. 177–204. The communication effect has also been observed in games involving 'minimal contributing coalitions' in which a few subjects, selected by a lottery, are asked to give up their endowments. All members of the group gain if the selected subjects contribute, but there is no mechanism to force them to do so once selected, and no way to reveal the identity of defectors. Nevertheless, Van de Kragt, Orbell, and Dawes found that if communication was allowed prior to selection, all the subjects selected to contribute did in fact contribute. Braver similarly found that 71 percent surrendered their endowment after group discussion. He suggests that agents perceived the outcomes of group discussion as a 'social contract', and felt committed to fulfill its terms. See A.J.C. van de Kragt, J.M. Orbell and R.M. Dawes, 'The Minimal Contributing Set as a Solution to Public Goods Problems', *American Political Science Review* 77 (1983): 112–22; S.L. Braver, 'Social Contracts and the Provision of Public Goods', in *Social Dilemmas: Perspectives on Individuals and Groups*, edited by D.A. Schroeder (Westport, CT: Praeger, 1995), pp. 69–86.
16. Norbert L. Kerr, Jennifer Garst, Donna A. Lewandowski and Susan E. Harris, 'That Still, Small Voice: Commitment to Cooperate as an Internalized Versus a Social Norm', *Personality and Social Psychology Bulletin* 23 (1997): 1300–11. The step-level social dilemma is one in which, after a threshold number of contributions is reached, the public good is provided. Such dilemmas involve a coordination element, since less than the full number of participants is needed to provide the public good.
17. Kerr et al., 'That Still, Small Voice: Commitment to Cooperate as an Internalized Versus a Social Norm'. Similar results were obtained by Jerdee and Rosen, 'Effects of Opportunity to Communicate and Visibility of Individual Decisions on Behavior in the Common Interest' and by Kerr and Kaufman-Gilliland, 'Communication, Commitment and Coordination in Social Dilemmas'.
18. Robyn M. Dawes, 'Social Dilemmas', *Annual Review of Psychology* 31 (1980): 169–93.
19. Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms*.
20. Iris Bohnet and Bruno S. Frey, 'The Sound of Silence in Prisoner's Dilemma and Dictator Games', *Journal of Economic Behavior and Organization* 38 (1999): 43–57; Brosig et al., 'The Effect of Communication Media on Cooperation'.

21. Kelly S. Bouas and S.S. Komorita, 'Group Discussion and Cooperation in Social Dilemmas', *Personality and Social Psychology Bulletin* 22 (1996): 1144–50. Similar results were obtained in an earlier study by Dawes et al., 'Behavior, Communication, and Assumptions about Other People's Behavior in a Commons Dilemma Situation'.
22. Numerous studies point to the strong normative content of discussions of social dilemmas. Discussants emphasize not only the mutual gains obtained from cooperation, but also its appropriateness and normative appeal. It has been argued that discussants aim simultaneously to establish a sense of self-esteem and dignity for cooperators and to provoke shame and guilt for not cooperating. Defectors are exposed to normative pressures, ranging from tongue-lashing and the expression of anger to implied (and incredible) threats about future sanctions that would follow defections, even outside the laboratory environment. In one experiment, Cason and Khan observed that 'Several participants made it a point to inform their fellows of the incentives to defect and of the importance of not defecting since, once trust had been destroyed, it would be impossible to cooperate again. Once a strategy had been agreed upon by the group, it was adhered to fairly well.' See Timothy N. Cason and Feisal U. Khan, 'A Laboratory Study of Voluntary Public Goods Provision with Imperfect Monitoring and Communication', *Journal of Development Economics* 58 (1999): 544–5. Likewise, see Norbert L. Kerr, 'Norms in Social Dilemmas', in *Social Dilemmas: Perspectives on Individuals and Groups*, edited by David A. Schroeder (Westport, CT: Praeger, 1995), pp. 29–47; Phillip Bonacich, 'Norms and Cohesion as Adaptive Responses to Political Conflict: An Experimental Study', *Sociometry* 35 (1972): 357–75; Phillip Bonacich, 'Secrecy and Solidarity', *Sociometry* 39 (1976): 200–8. See also Dawes et al., 'Behavior, Communication, and Assumptions about Other People's Behavior in a Commons Dilemma Situation'; Ostrom and Walker, 'Communication in a Commons: Cooperation Without External Enforcement'; Gardner et al., 'Social Capital and Cooperation: Communication, Bounded Rationality, and Behavioral Heuristics'.
23. Orbell et al., 'Explaining Discussion-Induced Cooperation'; Braver, 'Social Contracts and the Provision of Public Goods'.
24. For a much broader and formal discussion, see Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms*, Chs 1–2 especially. Several alternative hypotheses about the causes of the communication effect have been rejected in the past, due to carefully designed experiments that suggest they are insufficient to explain it. Such hypotheses indicate that communication may (1) provide information and facilitate understanding of the game, (2) promote coordination of cooperative action, (3) alter the expectations about other players' behaviors (even when promises are not involved), (4) invoke generic norms of cooperation, (5) humanize other players, or (6) contribute to the creation of group identity. For further analysis, see Robyn M. Dawes, Alphons J.C. van de Kragt and John M. Orbell, 'Cooperation for the Benefit of Us – Not Me, or My Conscience', in *Beyond Self-Interest*, edited by Jane J. Mansbridge (Chicago, IL: University of Chicago Press, 1990), pp. 97–110; Ostrom and Walker, 'Communication in a Commons: Cooperation Without External Enforcement'; Elinor Ostrom, '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 (1998): 1–22; Shankar and Pavitt, 'Resource and Public Goods Dilemmas: A New Issue for Communication Research'; Bicchieri, 'Covenants Without Swords: Group Identity, Norms and Communication in Social Dilemmas'.

25. Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms*, Ch. 1.
26. Sensitivity to the behavior of others has been demonstrated in many experiments. Schroeder et al. remark that in a dilemma experiment, subjects conformed to the behavior of the interacting others: 'the course of action that is behaviorally endorsed by others may be adopted as the "correct" strategy, regardless of the objective validity of this strategy'. See David A. Schroeder, Thomas D. Jensen, Andrew J. Reed, Debra K. Sullivan and Michael Schwab, 'The Actions of Others as Determinants of Behavior in Social Trap Situations', *Journal of Experimental Social Psychology* 19 (1983): 536. For similar findings, see also Scott T. Allison and Norbert L. Kerr, 'Group Correspondence Biases and the Provision of Public Goods', *Journal of Personality and Social Psychology* 66 (1994): 688–98; Madan M. Pillutla and Xiao-Ping Chen, 'Social Norms and Cooperation in Social Dilemmas: The Effects of Context and Feedback', *Organizational Behavior and Human Decision Processes* 78 (1999): 81–103; Cason and Khan, 'A Laboratory Study of Voluntary Public Goods Provision with Imperfect Monitoring and Communication', p. 546; Messick and Brewer, 'Solving Social Dilemmas'; Olivier Bochet and Louis Putterman, 'Not Just Babble: A Voluntary Contribution Experiment with Iterative Numerical Messages', URL (consulted 3 April 2006): www.personeel.unimaas.nl/o.bochet/www/notjustbabble.pdf.
27. Cialdini et al., 'A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior'.
28. Dawes observed that when the experimenter was directed to 'moralize at the subjects', they cooperated at rates similar to subjects who were allowed to communicate. See Dawes, 'Social Dilemmas', p. 188.
29. Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms*.
30. Scripts are cognitive structures that contain knowledge about events, roles, and so on (for example, a lecture, going to a restaurant, playing a chess game). They are like models insofar as they contain sets of stylized facts.
31. Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms*.
32. Lev-On, 'Web-Based Collaboration and the Organization of Democracy'.
33. This is likely to change in the next few years with further developments in experimental communication tools.
34. Brosig et al., 'The Effect of Communication Media on Cooperation'.
35. In the 'lecture' condition, subjects watched a videotape explaining the game, delivered by a lecturer who was not involved in the experiment. In the 'talk-show' condition, subjects watched a videotaped discussion about the experiment, carried out by the 'videoconference' group.
36. Frohlich and Oppenheimer, 'Some Consequences of E-mail vs. Face-to-Face Communication in Experiment'.
37. Olivier Bochet, Talbot Page and Louis Putterman, 'Communication and Punishment in Voluntary Contribution Experiments', *Journal of Economic Behavior and Organization* 60 (2006): 11–26.
38. Bochet and Putterman, 'Not Just Babble: A Voluntary Contribution Experiment

with Iterative Numerical Messages'. Subjects were asked to choose one of the following statements: 'I do not wish to make a promise at this time' or 'I promise to contribute to the group account this period.' If they chose the second option, they had to type an integer between zero and ten.

39. Nathan Bos, Darren Gergle, Judith S. Olson and Gary M. Olson, 'Being There versus Seeing There: Trust via Video', *Proceedings of CHI 2001* (2001): 291–2.
40. James Andreoni, 'Cooperation in Public-Goods Experiments: Kindness or Confusion?', *American Economic Review* 85 (1995): 891–904.
41. Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms*, Ch. 1.
42. Bos et al., 'Being There versus Seeing There: Trust via Video'; Elena Rocco, 'Trust Breaks Down in Electronic Contexts but can be Repaired by some Initial Face-to-Face Contact', *Proceedings of CHI 1998* (1998): 496–502.
43. Note that in the 'mailing list' condition, the communication medium was asynchronous (mailing list) and not synchronous (chat). This may explain the difficulties experienced by subjects in organizing a cooperative 'contract'.
44. In another public-goods experiment, Wilson and Sell also found a large discrepancy between promises and behavior: 53.4 percent of the pledges to cooperate resulted in contributions smaller than had been promised. See Rick K. Wilson and Jane Sell, 'Liar, Liar . . . Cheap Talk and Reputation in Repeated Public Goods Settings', *Journal of Conflict Resolution* 41 (1997): 695–717.
45. The contributions in the TTS condition (in which a computer 'reads aloud' the text entered by the user at the other end) were consistently (although not significantly) higher than in the text condition (in which the text entered by the users at the other end simply appeared on the screen of the other user). See Carlos Jensen, Shelly D. Fernham, Steven M. Drucker and Peter Kollock, 'The Effect of Communication Modality on Cooperation in Online Environments', *CHI Letters* 2 (2000): 1–6.
46. Zheng et al., 'Trust Without Touch: Jumpstarting Trust with Initial Social Activities'; Jun Zheng, Nathan Bos, Judith S. Olson and Gary M. Olson, 'Trust Without Touch: Jump-Start Trust with Social Chat', *Proceedings of CHI 2001* (2001): 293–4.
47. Bonacich observed in 'Secrecy and Solidarity' that institutionally and normatively 'thin' environments encourage greater normative exchange (that is, promises and threats) between agents, which results in stronger normative pressures. Frohlich and Oppenheimer found that the CMC condition in a public-good experiment was much more emotionally and normatively charged than a comparable FtF condition. See Frohlich and Oppenheimer, 'Some Consequences of E-mail vs. Face-to-Face Communication in Experiment', p. 401. For similar observations, see Rocco, 'Trust Breaks Down in Electronic Contexts but can be Repaired by some Initial Face-to-Face Contact'; Brosig et al., 'The Effect of Communication Media on Cooperation'.
48. Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms*.
49. David K. Lewis, *Convention: A Philosophical Study* (Cambridge, MA: Harvard University Press, 1969).
50. Herbert H. Clark and Susan E. Brennan, 'Grounding in Communication', in *Perspectives on Socially Shared Cognition*, edited by Lauren B. Resnick, John M. Levine and Stephanie D. Teasley (Washington, DC: American Psychological Association, 1991), pp. 127–49.
51. Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms*, Ch. 2.

politics, philosophy & economics

article

© SAGE Publications Ltd
London
Thousand Oaks, CA
and New Delhi

1470-594X
200706 6(2) 139–168

Computer-mediated communication and cooperation in social dilemmas: an experimental analysis

Cristina Bicchieri
University of Pennsylvania, USA

Azi Lev-On
University of Pennsylvania, USA

abstract One of the most consistent findings in experimental studies of social dilemmas is the positive influence of face-to-face communication on cooperation. The face-to-face 'communication effect' has been recently explained in terms of a 'focus theory of norms': successful communication focuses agents on pro-social norms, and induces preferences and expectations conducive to cooperation.[1] Many of the studies that point to a communication effect, however, do not further explore whether and to what extent the communication medium affects cooperative behavior. In this article, we ask if pro-social behavior can emerge and survive in computer-mediated environments. We show that, like face-to-face communication, computer-mediated communication also positively affects cooperation in social dilemmas, but cooperation is more difficult to establish and maintain. We argue that the discrepancy between the computer-mediated and the face-to-face communication effects is a consequence of the distinct capabilities of different media to focus agents on pro-social norms and to allow them to develop mutual expectations about future behavior.

keywords social dilemmas, laboratory experiments, communication, social norms, cooperation, computer-mediated communication

DOI: 10.1177/1470594X07077267

Cristina Bicchieri is Professor of Philosophy and Legal Studies, University of Pennsylvania, USA [email: cb36@sas.upenn.edu]
Azi Lev-On is Post-doctoral Fellow in the Politics, Philosophy, and Economics program, University of Pennsylvania, USA [email: el322@nyu.edu]

52. Ibid., Ch. 4.
53. Note that communication via videoconference supports much higher cooperation rates than via audio-conference. In an experiment which examined the influence of the richness of deliberative settings on group polarization, Sia et al. found that deliberation generated group polarization when it was conducted through an audio-conference system, but not when agents deliberated through a videoconferencing system. The phenomena studied in the Brosig et al. and the Sia et al. experiments are quite different, but their conclusions are similar: the use of an auditory channel alone for communication increases social distance among communicators much more than the use of a visual channel. See Choon-Ling Sia, Bernard C.Y. Tan and Kwok-Kee Wei, 'Group Polarization and Computer-Mediated Communication: Effects of Communication Cues, Social Presence, and Anonymity', *Information Systems Research* 13 (2002): 70–90; Brosig et al., 'The Effect of Communication Media on Cooperation'.
54. Bochet et al. claim that using a (synchronous) chat room for communication may be more conducive to cooperation than using asynchronous channels, such as email and newsgroups. This hypothesis still has to be empirically tested. See Bochet et al., 'Communication and Punishment in Voluntary Contribution Experiments'.
55. For similar conjectures, see Rocco, 'Trust Breaks Down in Electronic Contexts but can be Repaired by some Initial Face-to-Face Contact'; Frohlich and Oppenheimer, 'Some Consequences of E-mail vs. Face-to-Face Communication in Experiment'.
56. John M. Orbell, Alphons J. van de Kragt and Robyn M. Dawes, 'Covenants Without the Sword: The Role of Promises in Social Dilemma Circumstances', in *Social Norms and Economic Institutions*, edited by Kenneth J. Koford and Jeffrey B. Miller (Ann Arbor: University of Michigan Press, 1991), p. 129; Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms*. For similar findings, see Adam F. Simon and Heather E. Gorgura, 'Say the Magic Word: Effective Communication in Social Dilemmas', paper presented at the annual meeting of the American Political Science Association, Philadelphia, 2003 (a secondary analysis of data from Gardner et al., 'Social Capital and Cooperation: Communication, Bounded Rationality, and Behavioral Heuristics').
57. Bicchieri, 'Covenants Without Swords: Group Identity, Norms and Communication in Social Dilemmas'; Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms*, Ch. 4.
58. For conjectures in this vein, see Wilson and Sell, 'Liar, Liar . . . Cheap Talk and Reputation in Repeated Public Goods Settings'; Frohlich and Oppenheimer, 'Some Consequences of E-mail vs. Face-to-Face Communication in Experiment'; Elinor Ostrom, 'Collective Action and the Evolution of Social Norms', *Journal of Economic Perspectives* 14 (2000): 137–58.
59. Thomas R. Palfrey and Howard Rosenthal, 'Testing for the Effects of Cheap Talk', *Games and Economic Behavior* 3 (1991): 183–220; Bochet et al., 'Communication and Punishment in Voluntary Contribution Experiments'; Xiao-Ping Chen and S.S. Komorita, 'The Effects of Communication and Commitment in a Public Goods Social Dilemma', *Organizational Behavior and Human Decision Processes* 60 (1994): 367–86; Ostrom, 'A Behavioral Approach to the Rational Choice Theory of Collective Action – Presidential Address of the American Political Science Association 1997'. The experiments that Ostrom describes are the following:

Wilson and Sell, 'Liar, Liar . . . Cheap Talk and Reputation in Repeated Public Goods Settings', replicated a previous public-good experiment by Isaac and Walker, 'Communication and Free-Riding Behavior: The Voluntary Contribution Mechanism'; Rocco and Warglien replicated a previous common-pool resource experiment by Ostrom et al., 'Covenants With and Without a Sword: Self-Governance is Possible'. See Elena Rocco and Massimo Warglien, 'Computer Mediated Communication and the Emergence of "Electronic Opportunism"', Working paper 1996–2001 (Trento: Trento University, Department of Economics, 1996). For additional experimental results supporting the conclusions below, see Xiao-Ping Chen, 'The Group-Based Binding Pledge as a Solution to Public Goods Problems', *Organizational Behavior and Human Decision Processes* 66 (1996): 192–202; Robert Kurzban, Kevin McCabe, Vernon L. Smith and Bart J. Wilson, 'Incremental Commitment and Reciprocity in a Real Time Public Goods Game', *Personality and Social Psychology Bulletin* 27 (2001): 1662–73.

60. Orbell et al., 'Covenants Without the Sword: The Role of Promises in Social Dilemma Circumstances'; Bouas and Komorita, 'Group Discussion and Cooperation in Social Dilemmas'.
61. Sanford L. Braver and L.A. Wilson II, 'Choices in Social Dilemmas: Effects of Communication within Subgroups', *Journal of Conflict Resolution* 30 (1986): 51–62; Kinukawa et al., 'Partial Communication in a Voluntary-Contribution-Mechanism Experiment'.
62. Kinukawa et al., 'Partial Communication in a Voluntary-Contribution-Mechanism Experiment', p. 415.
63. Bochet and Putterman, 'Not Just Babble: A Voluntary Contribution Experiment with Iterative Numerical Messages'.
64. Bochet et al., 'Communication and Punishment in Voluntary Contribution Experiments'.
65. Bohnet and Frey, 'The Sound of Silence in Prisoner's Dilemma and Dictator Games'.
66. Bornstein, 'The Free Rider Problem in Intergroup Conflicts over Step-Level and Continuous Public Goods'.
67. Bos et al., 'Being There versus Seeing There: Trust via Video'.
68. Bouas and Komorita, 'Group Discussion and Cooperation in Social Dilemmas'.
69. Braver, 'Social Contracts and the Provision of Public Goods'.
70. Braver and Wilson, 'Choices in Social Dilemmas: Effects of Communication within Subgroups'.
71. Brosig et al., 'The Effect of Communication Media on Cooperation'.
72. Cason and Khan, 'A Laboratory Study of Voluntary Public Goods Provision with Imperfect Monitoring and Communication'.
73. Ibid., p. 539.
74. Chen, 'The Group-Based Binding Pledge as a Solution to Public Goods Problems'.
75. Chen and Komorita, 'The Effects of Communication and Commitment in a Public Goods Social Dilemma'.
76. Dawes et al., 'Behavior, Communication, and Assumptions about Other People's Behavior in a Commons Dilemma Situation'.
77. Frohlich and Oppenheimer, 'Some Consequences of E-mail vs. Face-to-Face Communication in Experiment'.

78. Hackett et al., 'The Role of Communication in Resolving Commons Dilemmas: Experimental Evidence with Heterogeneous Appropriators'.
79. Isaac and Walker, 'Communication and Free-Riding Behavior: The Voluntary Contribution Mechanism'.
80. Jensen et al., 'The Effect of Communication Modality on Cooperation in Online Environments'.
81. Jerdee and Rosen, 'Effects of Opportunity to Communicate and Visibility of Individual Decisions on Behavior in the Common Interest'.
82. Kerr and Kaufman-Gilliland, 'Communication, Commitment and Coordination in Social Dilemmas'.
83. Kerr et al., 'That Still, Small Voice: Commitment to Cooperate as an Internalized Versus a Social Norm'.
84. Kiesler et al., 'A Prisoner's Dilemma Experiment on Cooperation with People and Human-Like Computers'.
85. Kinukawa et al., 'Partial Communication in a Voluntary-Contribution-Mechanism Experiment'.
86. Kurzban et al., 'Incremental Commitment and Reciprocity in a Real Time Public Goods Game'.
87. Orbell et al., 'Explaining Discussion-Induced Cooperation'.
88. Ostrom and Walker, 'Communication in a Commons: Cooperation Without External Enforcement'.
89. Ostrom et al., 'Covenants With and Without a Sword: Self-Governance is Possible'.
90. Palfrey and Rosenthal, 'Testing for the Effects of Cheap Talk'.
91. Rocco, 'Trust Breaks Down in Electronic Contexts but can be Repaired by some Initial Face-to-Face Contact'; Rocco and Warglien, 'Computer Mediated Communication and the Emergence of "Electronic Opportunism"'.
92. Schmitt et al., 'Collective Action with Incomplete Commitment: Experimental Evidence'.
93. Van de Kragt et al., 'The Minimal Contributing Set as a Solution to Public Goods Problems'.
94. Van de Kragt et al., 'Doing Well and Doing Good as Ways of Resolving Social Dilemmas'.
95. Wilson and Sell, 'Liar, Liar . . . Cheap Talk and Reputation in Repeated Public Goods Settings'.
96. Zheng et al., 'Trust Without Touch: Jump-Start Trust with Social Chat'; Zheng et al., 'Trust Without Touch: Jumpstarting Trust with Initial Social Activities'.