The Effect of Communication Modality on Cooperation in Online Environments

Carlos Jensen
carlosj@cc.gatech.edu

Shelly D. Farnham
a-shellf@microsoft.com

Steven M. Drucker
sdrucker@microsoft.com

Peter Kollock
kollock@ucla.edu

ABSTRACT

One of the most robust findings in the sociological literature is the positive effect of communication on cooperation and trust. When individuals are able to communicate, cooperation increases significantly. How does the choice of communication modality influence this effect? We adapt the social dilemma research paradigm to quantitatively analyze different modes of communication. Using this method, we compare four forms of communication: no communication, text-chat, text-to-speech, and voice. We found statistically significant differences between the various forms of communication, with the voice condition resulting in the highest levels of cooperation. Our results highlight the importance of striving towards the use of more advanced forms of communication in online environments, especially where trust and cooperation are essential. In addition, our research demonstrates the applicability of the social dilemma paradigm in testing the extent to which communication modalities promote the development of trust and cooperation.

Keywords

Computer Mediated Communication, Online Interaction, Social Interfaces, Collaboration, Social Dilemma, CSCW.

INTRODUCTION

One of the most consistent and robust findings in the sociological literature is the positive effect communication has on cooperation and trust [8]. When individuals are able to communicate, cooperation increases significantly. There are however many open questions. How does the choice of communication modality affect this finding, and how significant are the differences between different forms of communication? This is an important question for designers of collaborative online environments.

The research area we draw upon is the multi-disciplinary work on *social dilemmas*. Social dilemmas are situations in which a reasonable decision on an individual level leads to collective disaster, that is, a situation in which everyone is worse off than they might have been otherwise. Models of social dilemmas capture this tension between individual and collective outcomes, and can therefore be used as a very powerful and broadly applicable probe to assess the level of cooperation and trust in a group. Since the 1950s, a large research literature has developed in this area (for reviews, see [8,9,10]).

There has been surprisingly little work that has applied social dilemma models to online interaction. Further, despite the knowledge of the positive effects of communication on levels of cooperation, little is known about what particular aspects of communication are responsible for these effects. Previous work has touched on some of these issues, including a recent paper by Rocco examining the difference between electronic and face-to-face communication [11].

There is a large literature of work examining several modes of communication and their relative effects on task performance [4, 5, 14, 15], especially in the CSCW and video-conferencing domains. However, these studies have not specifically examined the effect of media on the development of trust and cooperation.

In most studies, communication has been used as a general term, without attempting to distinguish the various effects of different modalities. A recent survey [13] argued that the salutary effects of communication are largely limited to verbal discussions. Written communication was not found to have a significant and consistent effect on cooperation levels in groups. However, the evidence in this review was indirect, and none of the studies discussed attempted to actually compare different communication modalities against each other.

Our hypothesis was that more immediate forms of communication (such as face-to-face or voice) would prove more effective in promoting cooperation than less immediate forms such as text chat. The sociological literature supports this hypothesis, as do casual observations. However, bandwidth limitations and new technologies have created a multitude of communication forms that are harder to evaluate. Is low quality video more conducive to cooperation than high-quality still images, and how do either compare to the use of 2D or 3D avatars? If any of the above were available with text-chat or text-to-

speech (TTS), would the combination be more effective than voice communication alone?

Given a specific domain or application, the most appropriate form of communication can often be determined through traditional user testing, looking at such variables as task performance or user-preference. Choosing the right communication modality is crucial, not only out of technical considerations such as bandwidth, but also to encourage and support the desired activities. ICQ[*] and NetMeeting[†] are both popular Internet communication tools. These two tools are designed for different audiences, and the choice of communication modality is an integral part of that design decision. ICQ supports lightweight, informal communication through a more anonymous and non-invasive text channel. NetMeeting on the other hand supports a more intimate form of communication through the use of voice and video. This work may be most directly applicable to the domain of electronic commerce because of its reliance on trust and cooperation between participants with no previous relationship.

The unique contribution of this study is to examine four computer mediated communication modalities in a carefully controlled experimental setting, an online environment, and to use quantitative models of cooperation and competition that are based on several decades of research. We chose to look at the most common forms of online interaction: no communication, text-chat and voice. We also included TTS communication for two main reasons: to test our method on a modality that was harder to classify than the others; and to provide some interesting insight into the differences between text and voice communication

The results of this study are important both theoretically and practically. Theoretically, this investigation helps to explain the fundamental dynamics of cooperation in online settings. Practically, it can also lead to concrete recommendations on the introduction of different communication modalities or on how limited bandwidth should be best used in order to encourage trust and cooperation.

METHODOLOGY

System Design

For this study we used a continuous, iterated, dyadic (two-person) Prisoner's Dilemma. The continuous nature of the game allowed for degrees of cooperation or defection to take place. By making the game iterative, relationships were allowed to evolve over time, negotiations to take place, and trends to develop.

Pre-existing relationships, no matter how brief, greatly influence future interactions, as demonstrated by the work done by Rocco [11]. Therefore, we decided to focus our efforts on studying the development of relationships, rather than how they change with the use of different mediums. We decided to pursue dyadic trials as opposed to larger groups for logistic reasons, organizing trials with two subjects is much simpler than doing so with four or five.

The Prisoner's Dilemma is the most commonly used 2-person social dilemma. The game captures the key tensions between individual and collective outcomes: There is a strong temptation to behave selfishly and exploit the partner, but both persons are hurt if they behave selfishly. What defines the Prisoner's Dilemma is the relative value of the four outcomes (see Table 1). The best possible outcome is defecting while the other player cooperates (termed DC). The next best outcome is mutual cooperation (CC) followed by mutual defection (DD). The worst outcome is the case when one cooperates while the other player defects (CD). Thus in a Prisoner's Dilemma: DC>CC>DD>CD.

		II	
		C	D
I	C	2 **2**	3 **0**
	D	0 **3**	1 **1**

Table 1: Prisoner's Dilemma

I *and* II *designate Player* I *and* II. C *and* D *designate cooperation and defection respectively. Player* I*'s outcomes are in bold.*

During the course of the game, subjects were allowed to communicate with each other using one of four forms of communication; no communication, text-chat, TTS and voice (via speakerphone). All other factors were kept constant across the four cases. The game itself was built on top of an IRC-like communication channel. All contributions were sent using hidden messages and intercepted by a controller-bot. This bot kept track of turn numbers, running totals and grouped the players. In the text chat condition, the client was divided into two parts, one housing the game itself, and the other housing a standard chat interface, including a history of the subject's conversation.

TTS was implemented through the publicly available Microsoft Speech API 4.0[‡]. The voice used in all cases was the default, slightly feminine voice. While gender has been shown to produce biases, it was judged to be the most understandable voice by pre-test subjects. The interface for the TTS case was similar to that of the text case, except that the text messages and the history were hidden from the user. This was done to force the user to rely exclusively on the TTS technology. The subjects also had their own

---

[*] Available at http://www.icq.com

[†] Available at http://www.microsoft.com/windows/netmeeting/

[‡] Available for from http://www.microsoft.com/iit/

messages read back to them in order to inspire greater confidence in the system. For the voice case, we had a speakerphone system in place and used only the game portion of the interface.

Game Rules

The particular type of Prisoner's Dilemma we used was a continuous version of the game in which *degrees* of cooperation were possible: Each turn the two players were allocated 10 points and given a choice of how many points (from 0-10) they wished to contribute to their partner. Any contributed points were doubled and given to the partner.

Therefore the situation had the structure of a Prisoner's Dilemma: The greatest possible return comes from keeping all of one's points while the partner contributes all 10 points (DC=30 points - the 10 original points plus the 20 points from the partner's doubled contribution). However, if both actors followed this strategy each will end up with only 10 points (DD - having contributed none to each other) rather than the 20 points each would receive if they both contributed all their points (CC).

In order to promote a high level of motivation (and risk), the subject's compensation was tied to their final score. Those attaining near perfect cooperation or those who consistently convinced the other player to contribute more than they did themselves, earned a piece of software of their choice. Lower scores resulted in less valuable prizes, culminating at the lowest score levels with a Frisbee or pen. The subjects were informed of this fact before the game.

The two players could see the amount each person had contributed on the previous round, as well as that rounds' outcome for both players. To minimize the effect of end-game conditions the subjects were told that the game would last for approximately 120 rounds. All the games ended after 96 rounds, giving the players a warning to this effect on turn 95. This was done in order to test for the presence of an end-game condition without tainting the rest of the sequence.

Subjects

The subjects were paired randomly and assigned to one of the four communication categories. They did not meet each other before, during, or after the game. Great care was taken to present the other subject in as neutral a language as possible, avoiding all terms such as partner or opponent. This was done in order to preserve the tension between the players without pitting them against each other. For the same reasons, we did not place any restrictions on topics of conversation. Subjects were allowed to negotiate freely or discuss their private life if they so chose.

A total of 90 adult subjects played the game using one of the four communication modalities. Our subject population was very diverse, the average age being 40, with the min and max ages being 19 and 58 respectively. The

occupations of the subjects ranged from being retired, to engineers, police officers or students. Prior to running the experiments, the subjects completed a tutorial explaining the rules and how to use the program. As part of this tutorial they were asked to complete 4 questions in order to demonstrate their understanding of the game.

After the experiment, subjects filled out a questionnaire with standard questions to rate among other things their level of understanding of the rules and level of motivation. This was done in order to exclude the subjects who did not understand the game and those who did not take it seriously enough.

Two dyads were excluded because at least one member reported that they did not have a clear understanding of the game. Three other dyads were excluded when it became apparent from their communication that they had fundamentally misunderstood the game. All 3 of these pairs believed that the equation CD+DC>CC+CC to be true. Every other turn they would alternate defection and cooperation, believing the average of this strategy to be better than continued mutual cooperation. This elaborate strategy required tremendous amounts of communication and coordination between the two subjects, demonstrating full cooperation and trust in the other subject. However, this strategy resulted in a less than optimal score.

Four additional dyads were excluded because at least one member reported that they were not motivated to earn as many points as possible. Three additional dyads were excluded from the analysis because of their refusal to use the communication modality offered to them. We discuss this issue later in the paper

After excluding the invalid data, we were left with a total of 66 subjects, or 33 dyads: 9 dyads in the non-communication case, 9 dyads in the text chat case, 7 dyads in the TTS case, and 8 dyads in the voice case.

RESULTS

We expected that the form of communication would have a significant effect on the level of cooperation between the dyadic pairs. The more personal or intimate forms of communication were expected to be more conducive to the development of trust and cooperation. A dyadic pair exhibited the highest degree of cooperation (CC) by exchanging all 10 of their points every round. Table 2 and Figure 1 illustrate that form of communication does affect average contributions as expected, and an omnibus test of condition, using an analysis of variance (ANOVA), shows that type of communication had a statistically significant effect ($F(3, 29) = 3.42$, $p < .04$).

Mode of Communication	Dyad N	Mean Contribution	Standard Deviation
A. None	9	5.3^{cd}	4.2
B. Text Chat	9	6.4^{d}	3.5
C. Text-to-Speech	7	8.4^{ad}	1.4
D. Voice	8	9.4^{abc}	0.2

Table 2: Mean dyadic contribution as a function of the mode of communication. A mean of 10 would indicate perfect cooperation between the dyadic pairs. Mean subscripts (abcd) indicate which other means are at least marginally significantly different (p < .09).
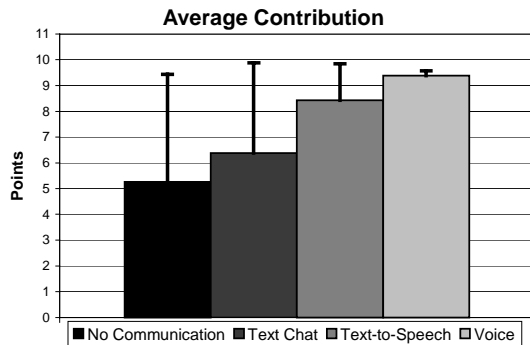

Figure 1: Mean dyadic contribution as a function of mode of communication. A mean of 10 would indicate perfect cooperation between the dyadic pairs. Error bars represent standard deviations.

Dyads in the voice condition showed the greatest levels of cooperation. Most dyads using voice communication exhibited almost perfect cooperation. Planned comparisons show that dyads in the voice condition on average contributed significantly more than dyads in the text chat condition ($F(1, 15) = 5.82$, $p < .03$) and the no communication condition ($F(1, 15) = 7.69$, $p < .02$). The difference in average contribution between the voice condition and the TTS condition was marginal ($F(1, 13) = 3.50$, $p < .09$).

Dyads also tended to be more cooperative in the TTS condition than in the text and the no communication condition. Dyads in the TTS condition contributed marginally significantly more than the dyads with no communication ($F(1, 14) = 3.66$, $p < .08$). The difference in contributions between TTS and text conditions was not significant ($F(1, 14) = 2.13$, $p < .17$), however the trend is in the predicted direction and could reach statistical significance with a larger number of dyads. That people showed greater levels of cooperation in the TTS condition than in the text conditions indicates that voice affects cooperation for reasons other than the differences in the semantic content of text versus speech, and for reasons other than the nonverbal information communicated through personal voice, such as intonation and gender.

While dyads in the text condition on average contributed more than dyads with no communication, the difference was not significant ($F(1, 16) = .37$, *ns*). Dyads showed a much higher variability in the levels of cooperation in the text condition and the no communication conditions than in the TTS or the voice conditions. A better understanding of the nature of the greater variability in the no communication and the text chat condition can be gained through an examination of histograms of the total score. See Figure 2. In the no communication condition, people tended to fall into either a pattern of no cooperation, or complete cooperation. People in the chat condition showed a range of levels of cooperation. In the TTS condition and the voice condition people tended to be at least somewhat cooperative.
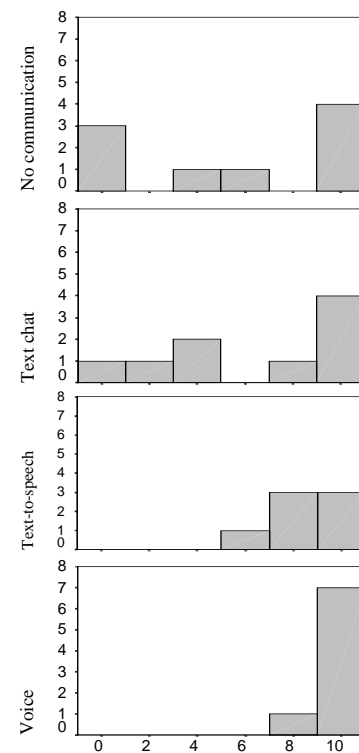

Figure 2: Distribution of average contributions for each mode of communication.

An analysis of contributions over time (see Figures 3 and 4) suggests that subjects did not initially cooperate fully with each other, but in conditions that allowed communication achieved a degree of cooperation that was greater than their initial level. A multivariate analysis of variance on 19 averaged blocks (see Figure 4), shows that there was a significant quadratic trend ($F(1, 29) = 5.14$, $p < .03$). An examination of Figure 4 indicates that the quadratic tend has a negative second-degree coefficient. However, the shape of the trend was not affected by condition ($F(3, 29) = .46$, *ns*).
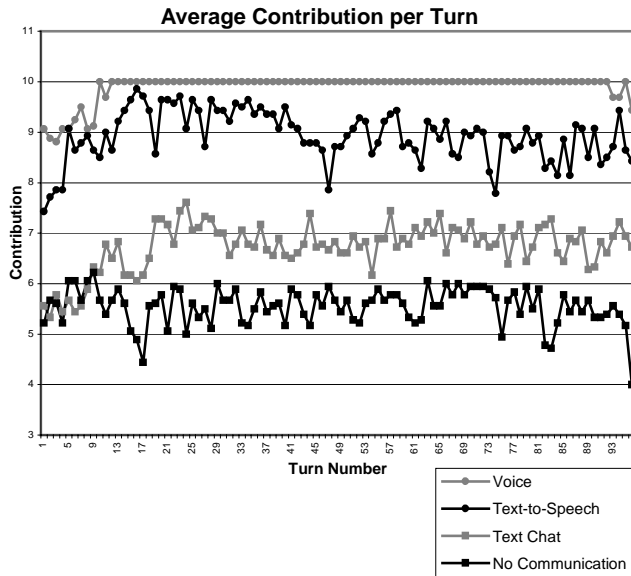
Figure 3: Average contribution across game rounds broken down by mode of communication.
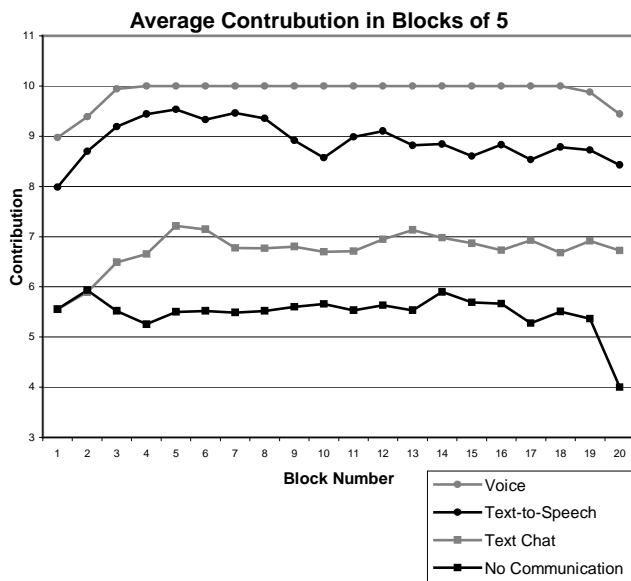

Figure 4: Average contribution across game rounds in blocks of five, broken down by mode of communication.

During the course of the game the subjects tended to vary their contributions across rounds (see Figure 3). The mean standard deviation for each dyad was 1.13, indicating that dyads tended to vary their contributions across rounds by adding or subtracting a point or two from their average.

We argued that people achieved greater levels of cooperation in the voice and text conditions because they were able to achieve greater levels of trust. To address whether people subjectively felt more trusting of the other across conditions, subjects were asked a series of questions in a post-experiment questionnaire. Subjects were asked to

rate the other player on characteristics related to trustworthiness (honest, fair, trustworthy, and sincere), likeability (likable, kind, friendly, and warm), and intelligence.

As seen in Figure 5, people felt more trusting and liking for the other player if they were able to communicate with that person. An ANOVA shows a marginally significant effect of condition on ratings of the other player's likeability ($F_{(3, 59)} = 2.44$, $p < .08$) and trustworthiness ($F_{(3, 59)} = 2.54$, $p < .07$). Only in the voice condition did people tended to rate the other player as more intelligent. There was a significant effect of communication modality on ratings of the other player's intelligence ($F_{(3, 61)} = 2.79$, $p < .05$)[§].

Figure 5: Evaluation of interaction partner

DISCUSSION

Consistent with sociological literature for off-line experiments, voice communication was found to have an extremely powerful effect on people's ability to trust and cooperate with each other. We found statistically significant differences between the voice communication condition and both the text-chat and the no-communication condition in predicting cooperation. We found marginally significant differences between voice and TTS, and between TTS and no-communication.

An examination of how people evaluated their partner suggest that they had a more positive image (likable and trustworthy) of those with whom they could communicate. In addition, people felt that their partners were more intelligent when they could communicate with them by voice. This effect could be caused by the differing amounts of time that it took for people to communicate in each of the modalities. Past studies have shown a direct relationship between delays in communication and perceived levels of intelligence [12].

One possible explanation for why we did not find statistically significant differences between TTS and text-chat is that our sample size was too small. A power analysis suggests that the effect could reach significance with approximately double the number of dyads (15 in each group as opposed to 8). Also consistent with previous findings in the sociological literature, we did not find statistically significant differences in cooperation between non-communication and text-communication modes. The lack of an effect may be due to the high variance in the two conditions. An examination of histograms of scores for each condition suggests that people tended to lapse into either full cooperation or full defection in the no

---

[§] This analysis was performed on ratings of the other player at the individual level rather than at the dyadic level. Dyadic effects were controlled for by entering the dyad as a covariate in the analysis.

communication condition, which greatly increases the variance of the condition.

Though our results are consistent with previous findings, we were nevertheless surprised by the lack of statistically significant differences between the text-chat and non-communication conditions. Strategic discussion was a primary topic of conversation in a number of the trials (approximately 80% of the dyads who could communicate discussed strategy), but while discussion was followed by robust cooperation in the voice condition, it was not as readily followed by cooperation in the text-chat condition.

We were also surprised at the relative difference between TTS communication and text-chat. Initially we expected no differences between TTS and text, because TTS contains the same semantic content as the text interaction, but little of the para-verbal information found in voice. In fact, since TTS might be artificially lacking cues that people use in predicting the veracity of utterances, we felt that there might be a negative effect of using TTS over text-chat interactions. There are several possible explanations for the higher level of cooperation in the TTS condition over the text-chat condition. Perhaps the experience of hearing a voice enhanced people's perceptions of social proximity. Anther possibility is that without an available history such as found in text-chat; subjects were forced to pay greater attention to the other player, hence increasing their sense of social presence.

The fact that TTS produced such high levels of cooperation compared to text suggests that it could be leveraged in other situations. Perhaps we could improve on voice or video interaction by modifying the voice or aspects of the appearance to match the preferences or biases of the other participant. For instance, instead of always using a gender-neutral voice, we could appeal to a user's sense of group identity by using the appropriately gendered TTS voice.

Finally, while the voice communication condition did achieve fairly universal cooperation amongst those who used this modality, a number of subjects showed some reluctance in actually using it. Many of these subjects reported being uncomfortable at the level of intimacy of voice communication, and said that they would have been more comfortable with text-chat. For these subjects, a less intimate mode of communication might well have been better than voice.

FUTURE WORK

In the future, we would like to run a larger number of subjects through these tests. With a larger number of subjects we believe we could show the benefit of text communication versus no communication, as well as see a clearer distinction between the other forms of communication. We would like to examine a number of additional modalities, including video at various resolutions and frame-rates, the use of both 2D and 3D avatars. As well

as asynchronous communication modalities such as voice mail and email.

The Prisoner's Dilemma can easily be extended to include a larger number of participants in every game. Not only do these N-person social dilemmas model an important class of interactions (group dynamics) not captured in the 2-person situation, they also create a significantly more competitive environment. This would help remove the ceiling effect of total cooperation we observed in the voice condition, and perhaps allow for a difference between voice and even more personal forms of communication (e.g. video and face-to-face) to be observed.

The social dilemmas are very sensitive to individual quirks and personality traits. To truly validate the results, a large number of subjects is needed, a costly requirement. By leveraging the World Wide Web (WWW), this cost could be significantly reduced. Such an approach would also be valuable for doing cross-national experiments. We are currently working on the design of such a system together with sociologists from a number of universities around the world.

Despite the simplicity of the task, a number of subjects were unable to figure out what the optimal game strategy was (CC). These subjects cooperated to the best of their abilities, but did so in a sub-optimal fashion (most of these believing DC+CD > CC+CC). Some subjects were thrown by the simplicity of the task, believing that there had to be a "trick", a secret strategy or hidden motivation for their task. In addition, many subjects remarked upon how boring the game was, and some stopped cooperating in an attempt to make things more interesting. We might be able to solve both problems by transforming the game from an abstract problem solving exercise into something more tangible, for instance some form of a card-game or video-game.

Further experiments need to be done on the sensitivity of the technique for comparing different communication modalities. We found that this method was very sensitive to factors such as wording and the perceived level of risk among the participants. If the stakes are low, people have a tendency to take higher risks and be less selfish. As the stakes get higher, the inclination towards selfish behavior increases. An interesting question is how to encourage a high sense of risk in a WWW version of the game.

We would like to start looking at factors other than cooperation, for instance the development of empathy, or how the medium affects your ability to determine identity. From our current results it seems that people might be prone to making assumptions for which they have no or little basis in the TTS case. In face-to-face interactions, many people believe that they can judge a person's honesty by simply looking into their eyes. Are people more prone to making such assumptions with some forms of communication than with others, and if so, how does the medium affect the assumptions that are made?

A final avenue of research is extending our results to pre-existing relationships. While there are many situations where we deal with relative strangers, the existing relationship addresses an important set of interactions. Pre-existing relationships have a very significant effect on future interactions, even in high-risk situations [11]. By studying how the medium affects existing relationships, our results could be more easily applied to CSCW and groupware applications.

CONCLUSION

We have demonstrated a technique for the quantitative assessment and comparison of the effect of different forms of communication on the development of trust and cooperation. Consistent with the sociological literature, voice communication was found to have an extremely powerful effect in fostering trust and cooperation. Also consistent with previous results, text chat was not found to have a statistically significant effect beyond that of no communication. In our experiments, this may be caused by the high variance of both the text and the non-communication cases.

As expected, more immediate forms of communication showed a greater impact on the development of cooperation. The biggest surprise of this study was the marked difference between TTS and text-chat. While there was not a statistically significant difference, the average contribution was 2 points higher in the TTS case. This indicates that TTS technology has the potential to positively influence computer-mediated communication.

The social dilemma methodology is a very robust and widely used tool in the social sciences. By leveraging this technique we are better able to analyze and evaluate the many factors affecting social interactions.

We believe that our methodology can be adapted to analyze a wider array of social factors. By going beyond trust and cooperation, our method could become a valuable general-purpose tool in the study of computer-mediated communication.

ACKNOWLEDGEMENTS

REFERENCES

1. Abela, A., and Sacconaghi, A., Value Exchange: The Secret of Building Customer Relationships On Line. *The McKinsey Quarterly 1997*, Number 2

2. Axelrod, R. *The Evolution of Cooperation*. Basic Books, New York, 1984

3. Chapanis, A., Ochsman, R. B., Parrish, R. N. & Weeks, G. D. Studies in interactive communication: The behavior of teams during co-operative problem solving. *Human Factors, 14.* (1972). 487-509.

4. Doherty-Sneedon, G., Anderson, A., O'Malley, C., Langton, S., Garrod, S., & Bruce, V. Face-to-Face and video mediated communication: A comparison of dialog structure and task performance, *Journal of Experimental Psychology: Applied, 3, 2,* (1997). 105-123.

5. Edigo, C. Teleconferencing as a technology to support co-operative work: Its possibilities and limitations. In J. Gallegher, R. E. Kraut, & C. Edigo (Eds.) *Intellectual teamwork: Social and technological foundations of cooperative work* (pp. 351-371). Hillsdale, NJ. Erlbaum Associates

6. Kelly, H. *Interpersonal Relations: A Theory of Interdependence*. Wiley and Sons, New York, 1978

7. Kiesler, S., Sproull, L. and Waters, K. A Prisoner's Dilemma Experiment on Cooperation With People and Human-Like Computers. *Journal of Personality and Social Psychology 1996*. Vol. 70, No 1, 47-65

8. Kollock, P. Social Dilemmas: The Anatomy of Cooperation. *Annual Review of Sociology 1998*. 24:183-214

9. Ledyard, J. Public Goods: A Survey of Experimental Research. In Kagel, J. and Roth, A. (Eds.), *The Handbook of Experimental Economics*. Princeton University Press, Princeton, NJ, 1995

10. Messick, D and Brewer, M. Solving Social Dilemmas. In Wheeler, L. and Shaver, P. (Eds.), *Review of Personality and Social Psychology*. Sage, Beverly Hills, CA, 1983

11. Rocco, E. Trust Breaks Down in Electronic Contexts but Can Be Repaired by Some Initial Face-to-Face Contact. *Proceedings of CHI '98* (Los Angeles CA, April 1998), ACM Press, 496-502.

12. Ruhleder, K., Jorden, B. Meaning-making across remote sites: how delays in transmission affect interaction. *Proceedings of 1999 European Conference on CSCW.*

13. Sally, D. Conversation and Cooperation in Social Dilemmas: A Meta-analysis of Experiments from 1958 to 1992. *Rationality and Society*, 7,58-92, 1995

14. Veinott, E. S., Olson, J. S., Olson, G. M., Fu, X. Video Helps Remote Work: Speakers Who Need to Negotiate Common Ground Benefit from Seeing Eachother. *Proceedings of CHI '99* (Pittsburgh PA, May 1999), ACM Press, 302-309.

15. Veinott, E. S., Olson, J. S., Olson, G. M., Fu, X. Video Matters! When communication is stressed video Helps.