The Researcher's Dilemma: Evaluating Trust in Computer-Mediated Communication

Jens Riegelsberger, M. Angela Sasse, John McCarthy

Address for correspondence
Jens Riegelsberger
j.riegelsberger@cs.ucl.ac.uk

Dept. of Computer Science
University College London
Gower Street
London WC1E 6BT, UK

The Researcher's Dilemma: Evaluating Trust in Computer-Mediated Communication

Jens Riegelsberger, M. Angela Sasse, John McCarthy
Dept. of Computer Science
University College London

Abstract

The aim of this paper is to establish a methodological foundation for Human-Computer Interaction (HCI) researchers aiming to assess trust between people interacting via computer-mediated communication (CMC) technology. The most popular experimental paradigm currently employed by HCI researchers are social dilemma games based on the *Prisoner's Dilemma* (PD), a technique originating from economics. HCI researchers employing this experimental paradigm currently interpret the rate of cooperation - measured in the form of collective pay-off - as the level of trust the technology allows its users to develop. We argue that this interpretation is problematic, since the game's synchronous nature models only very specific trust situations. Furthermore, experiments that are based on PD games cannot model the complexity of how trust is formed in the real world, since they neglect factors such as ability and benevolence. In conclusion, we recommend (a) means of improving social dilemma experiments by using asynchronous Trust Games, (b) collecting a broader range (in particular qualitative) data, and (c) increasing use of longitudinal studies.

1 Introduction

1.1 Trust in Computer-Mediated-Communications

Computer-mediated communication (CMC) technology supports interaction between people in many areas such as business and education today. Over the past decade, Human Computer Interaction (HCI) researchers have sought to identify specific usability criteria for these technologies. Within the traditional framework of task performance, user satisfaction, and user cost (Shackel, 1991), HCI researchers have investigated task effectiveness and efficiency (e.g. Olson, Olson, & Meader, 1995) and users' ratings of audio and video quality (e.g. Watson & Sasse, 1996) to decide whether technology can effectively support interaction. In recent years, the scope has been broadened to investigate more subtle and complex effects of the technology on users themselves (such as their physiological responses - Wilson & Sasse, 2000), their performance on highly complex tasks (such as the ability to detect deception – Horn, 2001) and interaction style or patterns (Monk & Watts, 2000; Vertegaal, Slagter, van der Holt, & Nijholt, 2001). The current focus of investigation is the degree to which users assess trustworthiness and establish trust in each other when interacting via CMC technology. Trust is an important consideration for two reasons:

1. Mediated interactions carry an increased risk. As users might be placed in different contexts or cultures, misunderstandings become more likely, and enforcement of agreements and regulations becomes more difficult. As risks increase and become more difficult to evaluate, users of collaborative technologies face more complex decisions. Trust helps to reduce this complexity – it is a shortcut for a full-scale, laborious evaluation of the risks

and benefits involved (Luhmann, 1979; Adams & Sasse, 2001; Lahno, 2002a). Hence, computer-mediated communication or interaction requires more *a priori* trust than face-to-face interaction.

2. Many users of advanced communication technology state that they find it hard to *develop* trust with someone they cannot see face-to-face. This problem is commonly attributed to the fact that these technologies do not convey the full richness of face-to-face encounters. They omit cues that are thought to be crucial for trust-building (Whittaker & O'Connaill, 1997; Döring, 1998; Mitra, 2002).

Thus, as communication technologies replace face-to-face encounters, there is the danger of a proliferation of low-trust interactions. In the long run, low-trust interactions are more costly than trust-based interactions due to the increased need for contractual agreements and external enforcements (Diekmann & Lindenberg, 2001; Handy, 1995). To evaluate the trust-building capability of different technologies, HCI researchers need valid and reliable measures of the trust users develop in their remote communication partners. Most CMC studies investigating trust employ social dilemma games based on the Prisoner's Dilemma as means of measuring trust within dyads or groups. These studies measure players' rate of cooperation and defection ('cheating') while they play these games over CMC-channels. The conditions studied include different media (such as text chat, audio, and video) and different forms of acquiring information about other players prior to distant interaction.

In this paper we critically review the current use of social dilemma games for trust measurement in CMC-research. First we define trust and introduce social dilemma games, with a special emphasis on games with a Prisoner's Dilemma (PD) structure, as these are widely used in CMC (section 2.1). We then consider their use as an experimental paradigm in economics and sociology to measure cooperation among individuals (section 2.2). In section 2.3, we specifically review current studies in CMC that employ social dilemma games to infer trust among users of distant interaction technologies. In section 3, we present the argument why the trust-building potential of a communication medium cannot be inferred from cooperation rates in PD-based social dilemma games alone. The main reasons are the synchronous nature of the games used (section 3.2), and the fact that social dilemma games model only limited aspects of trust (section 3.3).

1.2 Definitions

Trust is the chicken soup of social life. It brings us all sorts of good things – from a willingness to get involved in our communities to higher rates of economic growth (...), to making daily life more pleasant. Yet, like chicken soup, it appears to work somewhat mysteriously.
(Uslaner, 2002, p. 1)

The term *trust* is used in everyday language, but everyone has a slightly different understanding of what it actually entails. In the scientific community, the situation is unfortunately not very different. Trust has been studied for many years in many disciplines and there is a plethora of trust definitions researchers can choose from (Corritore, Kracher, & Wiedenbeck, 2001). The definitions contrast on different

dimensions and are applicable to different situations. In this paper, we focus on trust between individuals in organisational settings (Mayer, Davis, & Schoorman, 1995).

Trust is only required in the presence of *risk*. We need to trust when we lack exact knowledge of others' motives and abilities, and thus cannot foresee their future actions. Hence, the definition of trust most researchers (both in the domain of organisational studies and CMC) agree on is as follows.

> *Trust is the willingness to be vulnerable based on positive expectations about the actions of others[1].*

This definition is the lowest common denominator of trust. Corritore et al. (2001) call it an 'abstract definition', to distinguish it from operational definitions, which are applicable to specific real-world situations. The definition is abstract, because it does not name any sources of vulnerability nor does it clarify the structure of trust-requiring situations. As we will discuss later, trust situations differ widely in their structure (section 4.2) and in their sources of vulnerability (section 4.3). Both factors have a profound effect on the way trust is built. It is thus dangerous to take this abstract definition as a basis for experimental operationalisation without further discussing the effects of structure and source of vulnerability. The definition by Bacharach & Gambetta (2001) is more specific about the structure of situations requiring trust:

> *In general, we say that a person 'trusts someone to do X' if she acts on the expectation that he will do X when both know that two conditions obtain: if he fails to do X she would have done better to act otherwise, and her acting in the way she does gives him a selfish reason not to do X.*
> (Bacharach & Gambetta, 2001).

This definition differentiates between the *trustor* – the individual who acts first by trusting – and the *trustee* – the second mover who has to decide whether to cooperate or cheat. Hence it is a definition of *asynchronous* trust situations. Situations of mutual trust (when both individuals involved are trustor and trustee at the same time) are referred to as *synchronous* trust situations (see section 3.2). Furthermore, the above definition implicitly introduces the notion of *trustworthiness*. A trustworthy actor foregoes situational temptations based on incentives that are not defined by the situation at hand. Such incentives are based on *trust-warranting properties* of the trustee (Bacharach & Gambetta, 2001). Examples include norms, affection, or fear of punishment. We will discuss them in more depths in section 3.3.

2 Social Dilemmas

Situations in which individual rational outcome contravenes collective rational outcome are known as *social dilemmas*. In these situations, individuals' efforts to maximise their own outcome will result in a reduced outcome for everyone (including themselves). An example of such a situation is trading: each trader is better off

---

[1] This definition is used by Zand (1972), Boss (1978), Mayer et al. (1995), McAllister (1995), Rocco (1998), and Corritore et al. (2001).

receiving the other's good without giving his own. However, both traders acting in this way will result in no exchange of goods, which is less desirable for both than a completed exchange. Social dilemmas are often seen as prototypical situations that require trust: actors have to forego maximised situational benefit and they have to trust others to do the same in order to achieve a maximal collective outcome.

The *Prisoner's Dilemma* (Tucker, 1950; Flood, 1952) is the most prominent social dilemma, but not the only one. Other examples are *Assurance Games* and *Stag Hunt* (Kollock, 1998). Most experimental research based on social dilemmas uses an incentive structure that is based on the Prisoner Dilemma. Below we introduce this particular social dilemma in detail.

2.1 Prisoner's Dilemma

In experimental Prisoner's Dilemma (PD) games, two players' decisions are associated with monetary rewards. The outcome with the highest pay-off for one particular player is to defect while the other player cooperates (DC), next is cooperation by both players (CC), followed by joint defection (DD). The outcome with the smallest pay-off results from cooperating while the other player defects (CD). Table 1 exemplifies the pay-off in a PD game.

As can be seen from Table 1, the individually rational decision to defect will lead to suboptimal outcome for both players: Both will be left with an outcome of 1 rather than 2. Thus, if both players were fully rational and would only decide based on the pay-offs defined by the game, they would always choose to defect in a one-off game as this is the *Nash-Equilibrium* of the game (Nash, 1950). Independent from the other player's decision, this will always be the best decision in retrospective (Poundstone 1993). We refer to social dilemma games with a Prisoner Dilemma pay-off structure as PD-based games.

	Cooperate	Defect
Cooperate	2 2	3 0
Defect	0 3	1 1

Table 1: Pay-off distribution in a PD game. The row player's pay-offs are in the lower left corner of the cell. The values are for illustration only.

However, in experimental PD games and in real-world situations that have a PD structure, actors do not just act on situational pay-off, but on their *utilities*. Utilities are a reflection of the players' real preferences. These will be influenced by many factors other than just the rewards given in a one-off PD game. Experimental rewards as they are given in Table 1 are also called *endogenous pay-off*, as they are defined within a one-off game. These endogenous pay-offs will influence each player's utility, but the actual utility associated with each outcome will also be shaped by *exogenous pay-offs*. These are pay-offs associated by the individual to particular outcomes – and are governed by situational factors. Expectation of future interaction – in experiments often introduced through iterated games – transforms the utilities of the outcomes in an individual games. Conditional cooperation becomes a rational option as it might induce the other player to cooperate in future rounds (Axelrod, 1980; Dieckmann & Lindenberg, 2001). Examples of other pay-off transforming factors include desire to comply with norms, benevolence, or expected future

interaction beyond the experiment (see section 3). These arguments are captured in Table 2. It illustrates how the hierarchy of preferences might change when exogenous pay-offs are factored in. In the example given, the row-player attaches value to cooperation irrespective of the move made by the other player.

	Cooperate	Defect
Cooperate	2+*2*	0+*2*
Defect	3+*0*	1+*0*

Table 2: Hypothetical utilities of the row-player. In this example the row-player attaches value to cooperation irrespective of the other player's action.

Thus, experimental PD games are games of *incomplete information*, as a participant does not know the other participant's true utility for each outcome. The missing information about *exogenous pay-offs* must be estimated from other cues available. This conflict between endogenous and exogenous pay-off reflects the conflict between situational temptation and trust-warranting properties that are central to trust. This conflict makes PD games interesting models for trust research. The factors that lead people to cooperate in the face of conflicting endogenous pay-off are of major interest. Hundreds of studies have examined the conditions of cooperation in social dilemma situations by using experimental PD games.

2.2 Social Dilemma Games in Experimental Economics

The two-person PD game described above can be seen as a special case of n-person PD games, where a larger group of individuals interact with each other. In experimental economics, n-person PD games are seen as models of *public good problems* that are of core interest in macroeconomics and policy-making (Taylor, 1987). *Free riding* is an example for a public good dilemma: Each person is better off using the bus without paying, i.e. defecting, but if everyone does this, the service will not be provided. Clearly, for this example, real-world collaboration is enforced through control and penalties. One structurally defining element of these public good problems is that they can be seen as *synchronous* games. Due to the size of the group, an actor in a public good problem cannot base his or her decision on knowledge about all the other actors' decisions. A PD game's synchronous structure is thus a good model of such situations. Furthermore, public good dilemmas are *symmetric* dilemmas as all individuals are confronted with the same situational pay-offs. In public good studies, *collective pay-off* is interpreted as a measure of cooperation in a group. Joint pay-off will be highest if all players cooperate, and lowest for collective defection.
In the real world, individuals cooperate in the face of conflicting pay-offs. Raub & Weesie (2000) name three broad categories of exogenous factors that ensure cooperation in social dilemma situations: (1) *time* – expectation of future interaction; (2) *institutions* – e.g. law enforcement agencies; (3) *social networks* – e.g. by providing reputation information. These factors can transform the utilities associated with each outcome in a way that makes cooperation a rational choice. In experimental studies, researchers have identified further factors that influence cooperation, as they can exclude or control the effects of time, institutions, and social networks. Table 3 gives an overview on the six most important experimental treatments that influence cooperation in experimental games based on a prisoner dilemma structure. They have been identified by Sally (1995) in a meta-review of 130 treatments in 37 studies from 1958 to 1992.

POSITIVE EFFECT ON COOPERATION	SALLY'S VARIABLE NAME
Direct Communication (between players) Frequency of verbal communication between participants measured as ratio of discussion periods to decisions.	FREQDISC
Instructions The experimental instructions favoured cooperation over defection.	COOP
Indirect Communication (via experimenter) Experimenter elicited promises regarding their decisions in the game from participants.	PROMISES
Iterated Games Number of rounds the game was played with a fixed partner.	LNITER
NEGATIVE EFFECT ON COOPERATION	
Relative pay-off Temptation to defect - measured as the difference between defection and cooperation pay-off relative to the cooperation pay-off.	TEMP
Type of pay-off Real money used as an incentive.	MONEY
Instructions The experimental instructions asked participants to improve relative gain. A competitive orientation was induced.	COMP

Table 3: Treatments influencing cooperation in experimental PD games (out of 21 analysed by Sally (1995)).

In the context of CMC-research, the clear positive effect that verbal communication has on cooperation is most relevant. Researchers assume that face-to-face communication might help to create group identity or trigger internalised norms of social conduct (Poundstone, 1993; Sally, 1995; Frohlich & Oppenheimer, 1998). In a study relevant to trust in CMC, Frohlich & Oppenheimer (1998) compared the effect of email *vs.* face-to-face communication, and concluded that face-to-face communication is more effective in stimulating cooperation in PD games. Bohnet & Frey (1998) argue that face-to-face communication provides two independent factors that support cooperation – the information exchanged and the mutual identification achieved by looking at each other. They demonstrated that *silent identification* before the game also increases cooperation among participants, a result they attribute to a decrease in social distance.

2.3 Social Dilemma Games in CMC

2.3.1 Studies & Results

The effect of the communication channel is the core interest of social dilemma studies in CMC. *Joint pay-off*, the measure of cooperation, is interpreted as an indicator of trust the participants hold in each other. These studies commonly use social dilemma games with dyads (Jensen et al., 2000; Zheng et al., 2002; Davis et al., 2002) [2] or groups (Rocco, 1998; Bos et al., 2002). One study used one-off games, in which participants played only one round of a game with each other (Jensen et al.; 2000). The other studies relied on iterated games, in which the participants played several rounds of the game. In the iterated studies, communication between

---

[2] Davis et al. (2002) do not focus on cooperation but on rates of defection. They interpret this measure as an indicator of general 'bad behaviour'.

participants was allowed every five rounds of the game (Rocco, 1998; Bos et al., 2002; Zheng et al., 2002) or continuously (Davis et al., 2002).

EFFECT OF COMMUNICATION CHANNEL				
Author(s)	**Iterated?**	**Group?**	**Conditions**	**Further data**
Jensen, Farnham, Drucker, & Kollock (2000)	Yes	Dyad	Face-to-face, text-chat text to speech, phone conference	Rating of other participant: Likeability, Trustworthiness, Intelligence.
Bos, Olson, Gergle, Olson, & Wright (2002)	Yes	3 person group	Face-to-face, videoconference, phone conference, text chat	Post-experiment trust ratings of other players.
Davis, Farnham, & Jensen (2002)	Yes	Dyad	Text chat, text to speech, text chat & information sheet	
EFFECT OF MODE OF AQUAINTANCE				
Author(s)	**Iterated?**	**Group?**	**Conditions**	**Further data**
Rocco (1998)	Yes	Group	Face-to-face, email, pre-game face-to-face inter-action then email	Content analysis of messages Questionnaire on group identity
Zheng, Veinott, Bos, Olson, & Olson (2002)	Yes	Dyad	Face-to-face, social text chat, photo, information sheet, no communication	Post-experiment trust ratings of other players.

Table 4: CMC studies that used PD games to measure trust.

Most studies used games with continuous investments. These games - unlike the classic PD game, which only has binary decisions - allowed for cooperation or defection on a continuous scale. Rocco (1998), Bos et al. (2002) and Zheng et al. (2002) achieved this by framing the game as an investment game, called *Daytrader*. Davis et al. (2002) used a game with a binary PD game structure. This game was framed as computer game called *WindUp World*. Rather than just deciding on defection or cooperation, players navigated wind-up toys through a virtual world. When they met, the players had to decide whether they wanted to wind each other up (cooperation), or short-circuit the other player (defection). Table 4 gives an overview on the types of games used and on the conditions under research. These conditions can be divided in two sets. First, there are studies that compare the effect of the communication channel used while playing the game (Jensen et al., 2000; Bos et al., 2002; Davis et al., 2002). These studies conclude that richer media result in higher rates of cooperation. The *audio only* and *video and audio* conditions reach levels of cooperation that are similar to those reached in face-to-face communication. They are however reached after more iterations than in face-to-face interaction. Furthermore, cooperation is less stable without communication, as defection increases towards the end of games (Bos et al., 2002).

The second set of studies researched the effect of prior acquaintance and personal information on cooperation (Rocco, 1998; Zheng et al., 2002). The motivation for this approach is to investigate Handy's (1995) claim that trust *'requires*

touch'. Giddens (1990) calls this commonly observed grounding of distant interaction in face-to-face contacts *re-embedding*. Rocco (1998) did indeed find a positive effect of prior face-to-face acquaintance on cooperation. Zheng et al. (2002) researched whether this effect can be reached by other means. They isolated three aspects of face-to-face interaction: synchronous interaction, visual identification and exchange of personal information. They compared the effect of face-to-face interaction, social text-chat, a photograph of the participant, and a personal information sheet[3]. Davis et al. (2002) also included a personal information sheet in their study. The results from this set of studies draw a picture similar to the studies in the other set: The richer the channel by which participants got to know each other, the higher the cooperation rate in the social dilemma game. Both studies report a very limited effect of the personal information sheet.

2.3.2 Cooperation as a Measure of Trust

The interpretation of cooperation as trust in situations with a PD structure is based on the abstract definition of trust *as willingness to be vulnerable based on positive expectations about the actions of others* (see section 1.2). Indeed, the decision to cooperate carries more risk: A player will regret having chosen to cooperate fully if the other player chooses to defect completely. When choosing to defect, a player will never have to regret the choice, as it will always be the better option independent of the other player's decision.

However, we argue that PD-based social dilemma games can be too limited when assessing a technology's trust-building potential for two major reasons: (1) Social dilemma games based on a PD-structure model a very specific subset of trust-requiring situations where decisions on cooperation and defection have to be reached synchronously: in PD-based games the trustor is at the same time trustee. This introduces a further risk, *strategic insecurity,* and makes the interpretation of cooperation and defection difficult (see section 3.2). (2) Wilful defection is only one source of vulnerability. There are further ones that are difficult to model in social dilemma games (see section 3.3).

3 Critique of Social Dilemma Games in CMC-Research

3.1 Risk Modelling

Trust has been described as a device to reduce complexity, a shortcut to avoid complex decision-making processes when facing decisions that carry risk. Trust is of particular importance when an individual knows little about the type, quantity and gravity of outcomes associated with a particular decision. As stated in section 1, this is one of the reasons for the importance of trust in distant interaction. Users of communication technologies face more – and often less understood – risks than interactants in face-to-face situations (Riegelsberger & Sasse, 2001). In a social dilemma game, participants are very well aware of the gravity and nature of risks associated with their decisions. However, many real-world risks will be perceived as threats, rather than risks, as they cannot by quantified by the actors. Trust is needed

---

[3] This sheet contained the information commonly exchanged in a get-acquainted small talk.

precisely when individuals cannot 'calculate' a utility for every possible outcome, or where this calculation would be too costly (Kee & Knox, 1970).

Secondly, for ethical reasons, all experiments lack the ability to create serious risks for the participants. Participants in the studies do not lose money in the course of the experiment – rather, their decisions determine the size of their *gain*. From a mathematical point of view, there may be no difference between a loss, and missing a gain that was expected. However, Kahnemann, Slovic, & Tversky (1982) showed that individuals' decisions depend on whether the outcomes are framed as loss or gain: generally, people are keener to avoid loss than to achieve gain. As a further example, sharing an idea in a brainstorming exercise carries the risk of embarrassment if the idea was only half-baked or – on the other hand – the risk of not being acknowledged as having conceived a good idea if it is taken on board by the whole team. Thus, whereas real-world risks are usually multi-dimensional and include non-monetary cost (such as embarrassment, loss of privacy or cognitive cost), in a social dilemma experiment, risks are generally monetary.

Furthermore, in the real world, it is often impossible for trustors to find out whether they have been cheated, or whether they are victims of unlucky circumstances. A colleague who does not deliver an important milestone on a project might just have had to prioritise another task or she might have wilfully sabotaged the project. The *a priori* trust one has in that colleague will determine how their behaviour is interpreted, and how one will react. Zheng et al. (2002) addressed this problem by introducing 'market fluctuations' in their version of the *Daytrader* game. These market fluctuations allow participants to withhold part of their investment and then blame the fluctuations for a lower than expected joint pay-off. This measure makes it impossible for participants to clearly identify defection. The complexity of their considerations is thus increased.

3.2 Synchronous Social Dilemma Games

If I do not trust the other player in a PD game to cooperate, I will choose not to cooperate myself - which makes me a defector. Even though I might want to act in a trustworthy manner, I have to defect pre-emptively if I do not trust the other player, to avoid being the 'loser' (Poundstone, 1993; Bacharach & Gambetta, 2001). This also holds true for games that allow continuous investments, in particular those that include windfall payments (cf. Zheng et al., 2002; Bos et al., 2002): I may want to invest fully, however, as other players might not invest fully, I should pre-emptively withhold some of my investment. As the windfall payments amplify minor differences, they make withholding investment very attractive in the rounds prior to the payout. Thus, pre-emptive withholding investment also becomes very likely.

As each player in a PD game is trustor and trustee at the same time, defection might take place due to any of the following reasons:
1. *Egotism*: The player has an individualistic or competitive orientation (Gallo & McClintock, 1965). The player is not trustworthy (has no trust-warranting properties) and just acts on the situational pay-off.
2. *Lack of trust*: Defection takes place due to defensive motives, as the other player is expected to defect. The player might want to withdraw rather than defect, but the PD game forces him or her to act as a defector if he or she lacks trust.
3. *Expectation of not being trusted*: The fear that the other player will act based on motive 2, i.e. he will defect pre-emptively, could trigger defection.

This list could be continued infinitely, with each player trying to second-guess the motives of the other. It results in a further risk that is specific to synchronous situations, where all actors are trustor and trustee at the same time and have to decide concurrently. This additional risk is called *strategic insecurity* (Lahno, 2002a). It cannot be overcome by complete information about the other player's motives and preferences, but depends solely on the specific synchronous and symmetric structure of the game. In section 2.1, we have argued that this kind of strategic insecurity is an important hurdle to mutual trust in settings such as the *public good problem*. In organisational settings, however, even when we speak about *mutual trust*, we rarely refer to situations where two individuals have to reach a trust-decision at exactly the same time without knowing the other's decision (Lahno, 2002a). Trust is more likely to manifest itself and grow through sequential asynchronous trust-requiring situations, where on any given instance one individual is the trustor and the other individual the trustee (Luhmann, 1979; Bacharach & Gambetta, 2001).

3.3 Trust-Warranting Properties

Trust has been defined as an act that increases one's vulnerability (see section 1.2). This definition, however, leaves open what the source of this vulnerability could be. In a social dilemma situation, the other player's decision to defect is clearly a source of vulnerability – but there are others. Deutsch (1958; 1960) sees *motivation* and *ability* as the basis of trust. If the trustee lacks either, the trustor makes himself vulnerable by trusting. Motivation of the trustee ensures that he will not choose to exploit the trustor's vulnerability, and ability ensures that he will not involuntarily harm the trustor. In their work on organisational trust, Mayer et al. (1995) reviewed several definitions of trust that specify personal trust-warranting properties. They synthesise them into the categories of *ability*, *benevolence* and *integrity* (see Table 5). We add to these *incentive structures,* as trust-warranting properties that are not properties of an individual but an important real-world trust ensuring mechanism. In our view, social dilemma games do not model ability and only a subset of motivation as a source of vulnerability.

TRUST-WARRANTING PROPERTY		DEFINITION
ABILITY		"… group of skills, competencies, and characteristics that enable a party to have influence within some specific domain." (Mayer et al., 1995) **– specific to a domain**
MOTIVATION	Benevolence	"… the extent to which a trustee is believed to want to do good to *the trustor,* aside from egocentric profit motive." (Mayer et al., 1995) **– property of trustor-trustee relationship**
	Integrity	"… the trustee adheres to a set of principles that the trustor finds acceptable." (Mayer et al., 1995) **– property of trustee**
	External Incentive Structure	Organisational or institutional rules a trustee is subject to. (cf. Lahno, 2002b)

Table 5: Trust-Warranting Properties.

3.3.1 Ability

In virtual teams – a core user group of CMC technologies – an individual's wilful decision to defect or 'cheat' will not be the main source of vulnerability. Yet, in such a setting motivation is clearly of importance, as it ensures that the virtual team member is engaged in the task and does not prioritise others. However, if a critical task is assigned to a virtual team member, the concerns would not only include that she might not be *motivated* to perform, but also that she might not be *able* to perform, or that she misinterprets the instructions. Risks that are based on lack of competence or misunderstandings are more salient than wilful defection.

Ability is domain-specific. Abilities that pertain to different domains have to be signalled in different ways, and are likely to require different channels to manifest themselves. As an example, an assessment of ability will differ depending on whether the task at hand is to jointly prepare a presentation or to jointly write a piece of code. Any study incorporating ability as a base of trustworthiness will have to address this issue when generalising to other situations.

3.3.2 Integrity

Integrity – a property of the trustee – ensures motivation by internalised codes of conduct, norms and values. An evaluation of *integrity* is based on an evaluation of the norms and values an individual claims to act on and an assessment of the conformity of his or her actions with these claims. Iterated social dilemma games with fixed identities that allow communication do allow an assessment of integrity, as participants can communicate their values and their conformance can be observed. Thus, the games used in recent CMC studies allow for trust based on integrity.

3.3.3 Benevolence

Benevolence is an appreciation of the trustor's good that forms a non-monetary incentive (exogenous pay-off) for the trustee to act as promised. It is not a property of the trustee, but a property of the relationship between the trustor and trustee.

After initial empathy or liking is signalled, benevolence can be built through several risky interactions (McAllister, 1995). The initial signalling of empathy and liking can be very fast, and might to a great extent rely on immediate affective

reactions. Such immediate effects can be tested within the social dilemma paradigm (Riegelsberger, Sasse, & McCarthy, forthcoming). However, strong feelings of benevolence depend on enduring relationships, built through many iterations of giving and receiving trust. Thus, in our view, benevolence is only marginally captured by current social dilemma games. These considerations call for longitudinal studies that extend beyond single experimental sessions (see section 5.3).

3.3.4 External Incentive Structure

Another important motivation not to defect in real-world settings is an external incentive structure. The most common source of such an incentive structure are institutions (such as e.g. companies or law enforcement bodies) that establish and enforce rules to which individuals have to submit. Lahno (2002b) refers to trust based on association with an organisation as *institutional trust*. Organisations, through explicit rules and hierarchies, provide incentives for their members to act in specific ways. A trustor by trusting an organisation and the rules it is built on will thus transfer some of this trust to representatives of the organisation: *"She's with Accountants Unbribable, she must be trustworthy."* Organisations may thus act as implicit third-party trust providers. Drawing again on the example of a virtual team, a team-members motivation to perform might be ensured by organisational rules, hierarchies and control functions.

This important aspect of trust in CMC is not taken in account in current social dilemma based trust research. The channel used or the context in which interaction is placed conveys much information on organisational association that will influence the perception of trust. Being contacted by someone on the company's Intranet will result in a different level of *a priori* trust than being contacted via an Internet Relay Chat (IRC). Modelling this important aspect of trust transfer is difficult in a laboratory setting, as the institution that gives access to and provides the channel will be the researcher.

3.3.5 Cognitive & Affective Trust

Another high-level categorisation of trust is the distinction between *cognitive* and *affective* trust. This distinction is not based on properties of the trustee or his incentive structures, but on how cues of trustworthiness are processed and how trust is formed by the trustor. McAllister (1995) based his model of organisational trust on this distinction and Rocco et al. (2000) and Corritore et al. (2001) introduced it to trust research in CMC.

Cognitive trust has also been termed *reliance*, *assurance* or *confidence* (Lahno 2002b, Yamagishi & Yamagishi, 1994). A trustee can encourage cognitive trust by signalling *ability* (see section 3.4.1), internalised norms of conduct (*integrity)* or by showing that his actions are governed by an external incentive structure (see section 3.3.4). Such *external incentive structures*, an important source of cognitive trust, are often given in the form of institutional rules to which an agent submits. However, trust encompasses more than cognitive elements. Lahno (2002a, 2002b) defines trust as an emotional attitude, which encompasses cognitive as well as affective dimensions. This is mirrored in Mayer et al.'s (1995) definition that includes benevolence as one of the core dimensions of trust. We posit that affective trust on an initial level is based on interpersonal cues that result in immediate affective reactions (Goffman, 1956; Reeves & Nass, 1996). A medium's capability for immediately creating affective trust responses can be analysed with social dilemma games. Social dilemmas are very valuable tools to investigate immediate effects when e.g.

comparing different channels for communication. They allow us to investigate the immediate effects on affective trust isolated from estimates of ability or integrity and without the drawbacks of other methods that are subject to participants' post-rationalisation. However, affective trust goes deeper than just immediate affective reactions, to include benevolence. As discussed above, social dilemma games are very limited with regard to including this trust-warranting property.

Affective trust is an important dimension of trust that should not be underestimated in organisational settings, as it supports self-organisation (coordination) of work teams and reduces the need for costly formal control structures (Rocco et al., 2000; McAllister, 1995). Figure 1 gives an overview of the dimensions of trust that are thought be of importance in different distant communication settings.
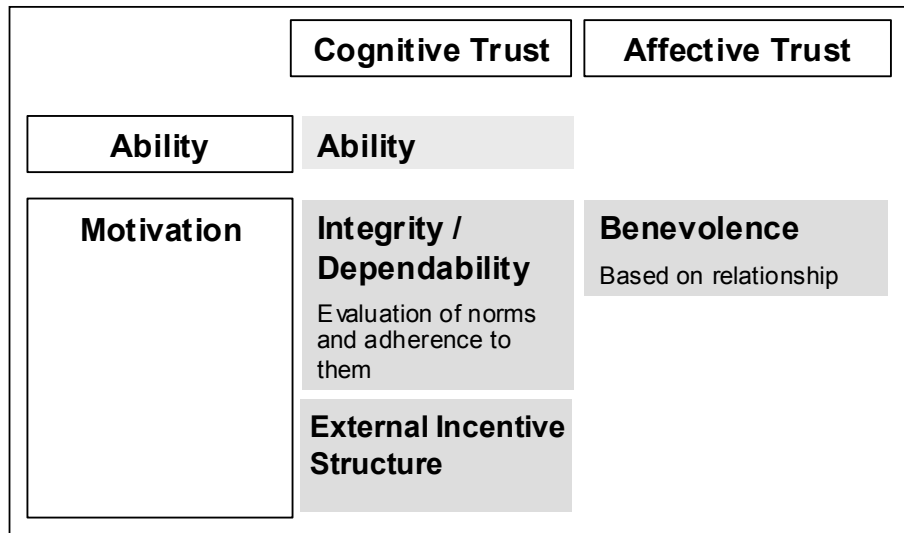

Figure 1: Bases of trust and their relationship with cognitive and affective trust.

4 Discussion

In section 3 we introduced several problems associated with PD-based social dilemma games as they are commonly used in CMC. In this section, we suggest improvements within the game-theoretical paradigm, and where appropriate advocate the use of other methods.

4.1 Taxonomy of Trust

We showed that real-world trust situations vary widely in terms of the type and gravity of *risk* they pose. Furthermore, one of the very reasons for trusting is that risks are not clearly identified or quantified and thus take the form of potential threats. A second dimension on which trust-requiring situations vary, and where PD based games only model a very small subset, is the *structure* of such situations. As discussed in section 3.2, PD-based games are models of synchronous trust situations. This results in an additional specific risk called *strategic insecurity*. Most real-world situations are not synchronous, and thus can be better modelled by another class of social dilemma games (see section 4.3). In our view, adequately modelling *sources of vulnerability* other than wilful defection is the main problem within social dilemma games. In a negotiation task, an assessment of integrity might be the main building block of trust, whereas in a virtual team setting, ability will be of high importance when assigning important tasks.

We identified three dimensions in which trust-requiring situations differ: *risks*, *structure*, and *sources of vulnerability*. The experimental approach taken to predict outcome in real-world situations should be determined by a situation's location in this space. Thus, as a first step we see the need for a taxonomy of situations requiring trust that incorporates these and other relevant dimensions.

4.2 Improved Risk Modelling

Social dilemma games used as models for trust situations do lack real risks, i.e. the possibility of loss. Creating *surrogate loss* can overcome this problem. Based on a study by Yates & Watts (1975), researchers can create loss in experimental settings by handing out rewards that then can be lost in the course of the experiment. Yates & Watts (1975) maintain that participants act on this surrogate loss in a similar way to real loss. Another problem we identified is that unlike in real-world settings, defection is clearly identifiable in social dilemma games. 'Market fluctuations' as they are used by Zheng et al. (2002; see section 3.1) increase the complexity of the participants' considerations. Both approaches are, however, only partial improvements, as they still only address the dimension of monetary pay-off. In real-world settings, risks and benefits are likely to span several dimensions, such as embarrassment, pride or cognitive effort. Experimental trust research is bound by a lack of real and complex risk.

4.3 Improved Social Dilemma Games

4.3.1 Trust Games

As real-world situations are unlikely to be synchronous and symmetric trust situations as modelled by PD-based games, experimental social dilemma games should be based on *asynchronous Trust Games*. Several researchers used asynchronous and asymmetric social dilemma games (Dasgupta, 1988; Kreps, 1990; Berg, Dickhaut, & McCabe, 1995; Hayashi, Ostrom, Walker, & Yamagishi, 1999; Bacharach & Gambetta, 2001). We introduce the *Trust Game*, as described by Bacharach & Gambetta (2001): Like the PD game, it opposes collective and individual rationality. However, unlike in the symmetric PD game, the participants act on different pay-off tables. In a Trust Game, the trustor first decides whether to trust the trustee or not. Thus, decisions are taken asynchronously. The decision not to trust, results in withdrawing from the interaction. The hierarchy of outcomes for the trustor is as follows: (CC) trustee cooperates > no interaction > (CD) trustee defects[4]. If the trustor decides to interact, the trustee can then decide whether to cooperate or defect: His experimental (endogenous) preferences are those of a PD game: (CD) defection > (CC) honouring given trust > no interaction. These make the situation risky for the trustor. Table 6 exemplifies the pay-off in a Trust Game. *Trust-warranting properties* (see section 3.3) may transform this endogenous hierarchy of preferences into: (CC) cooperation > (CD) defection.

---

[4] This is the hierarchy of preferences as it is given in Bacharach & Gambetta's Trust Game. However, there are real-world situations, where being defected is preferable to withdrawal.

Trustee ⇒ / ⇓ Trustor	Cooperate	Defect
Cooperate	1 3	4 -3
Withdraw	.. 1	.. 1

Table 6: Example of a Trust Game pay-off table (based on Bacharach & Gambetta (2001)).

This conflict between situational incentives to defect (based on experimental pay-off), and exogenous factors based on trust-warranting properties, is the very definition of a situation where trust is required (see section 1.2). Asynchronous trust games have increasingly received attention in economics and sociology, both in the form of experimental studies and in game-theoretical modelling (e.g. Raub 1992; Snijders 1993; Gautschi, 1999). They have also been used to test the effect of different reputation sharing mechanisms on giving and honouring trust (Bohnet, Huck, & Tyran, forthcoming). Bohnet et al. (forthcoming) framed their version of the Trust Game as a sale: The trustor decides whether she wants to buy from a seller or not. If the trustor has decided to buy, i.e. risk an amount of money, the trustee can decide whether he wants to fulfil or defect. A continuous version of the Trust Game allows the trustor to send an amount of money to the trustee. Any amount sent is tripled by the researcher. The trustee then can decide how much to send back to the trustor (Berg, Dickhaut, & McCabe, 1995). In this example, any trustor purely acting on endogenous pay-off would send nothing, as she would expect the equally endogenously motivated trustor to return nothing.

Thus, Trust Games allow for a separate analysis of the effects of technology on trusting and acting trustworthy. This is highly relevant to HCI researchers and practitioners, who should be concerned about designing technologies that allow for correct trust attribution, rather than just elevating users' trust levels. A further benefit of Trust Games is that they do not create *strategic insecurity* a specific risk that is associated with symmetric and synchronous trust situations, such as public good dilemmas (see section 3.2). As participants in PD-based games are trustor and trustee at the same time, they have to base their decisions not only on an evaluation of the other player, but also on how the other player might evaluate them (see section 3.2).

In summary, using asynchronous Trust Games, rather than synchronous PD-based games, allows for analysing giving trust and honouring trust separately, free participants from recursive considerations and model many real-world situations – that are mostly asynchronous – more accurately.

4.3.2 Symbols and Symptoms

A further advantage of Trust Games is that they allow modelling the problem of trust as a *signalling problem* (Bacharach & Gambetta, 2001). In order to be trusted, the trustee needs to signal to the trustor, that he has 'pay-off transforming' properties; that his actions are also governed by exogenous payoffs – brought about by e.g. specific norms, moral values, benevolence or care for own reputation rather than by the pay-off given in the current situation. Someone who can signal the possession of these properties is said to appear trustworthy. However, the situation is further complicated as individuals acting on raw payoff (i.e. non-trustworthy individuals) will also aim to appear trustworthy, so they will aim to send the signals that are associated

with the possession of these trust-warranting properties. A problem with current communication technologies is that – in the perception of users – they make the emulation of these signals cheaper when compared to face-to-face communication (Riegelsberger & Sasse, 2001). It is easier to put on a friendly smile on a photograph that forms part of an online profile than it is to act friendly in a face-to-face encounter. The same holds true for expressing benevolence or adherence to trust-ensuring norms. Due to its richness, face-to-face interaction makes it hard to control all the cues given off. Less rich media, due to their lack of cues, make it easier to control these fewer cues that are transmitted (Riegelsberger & Sasse, 2001). Hence, signifiers that can be treated as symptoms of trustworthiness (i.e. they are robustly associated with trust-warranting properties) become mere symbols when mediated and thus they do not allow users to infer the trustee's trustworthiness[5].

Hence, when designing technology to enable correct trust attribution, we must make it expensive for non-trustworthy individuals to emit the signals associated with trustworthiness, and cheap for trustworthy individuals to signal their intention or the possession of required trust-warranting properties.

4.4 Trust-Warranting Properties

The main disadvantage of social dilemma games in trust research is that they do not address sources of vulnerability adequately. Firstly, as discussed in section 3.3, the effect of some-trust warranting properties, such as ability, are hard to model. Secondly, they only allow researchers to observe the outcome of decisions, rather than giving indications as to what reasoning they are based on.

Ideally, Trust Games should be coupled with pre-and post-experimental interviews to clarify the participant's exogenous motivations. Qualitative analysis of verbal data – using techniques such as Grounded Theory (Strauss & Corbin, 1998) and Discourse Analysis (Potter & Wetherell, 1987) – have been successfully employed to elicit user perceptions and attitudes on complex issues such as privacy (Adams & Sasse, 2000), security behaviour (Adams & Sasse, 1999; Weirich & Sasse, 2001), and trust in e-commerce vendors (Riegelsberger & Sasse, 2001). In CMC trust research, they could also be used to establish participants' perceptions of each other in a more complex way than rating scales can and thus give insight into the users' basis for decisions.

These approaches carry, however, all the known weaknesses of self-reported data – such as *post-hoc rationalisation* (Norman, 1993). This problem is particularly virulent in trust research, as trust is based not only on cognitive, but also on affective dimensions (see section 3.3.5). A potential remedy to this problem is employing a *constructive interaction scenario* (Miyake, 1986), which HCI researchers have successfully employed to elicit users' models in ecologically valid settings (e.g. Clark & Sasse 1997). An experimental setting would pair up participants in such a way that they have to reach a joint decision, and thus have to verbalise their trust perceptions in a natural setting. To gain a more complete picture of the impact of interaction on the participants, and pinpoint critical phases during the interaction, objective physiological measures of user cost could be employed (see Wilson & Sasse, 2000).

---

[5] The notion of symbols and symptoms of trustworthiness was developed during a seminar at the University of Duesseldorf, Department of Sociology, organised by Bernd Lahno and Michael Baurmann.

5 Conclusion

The question of how HCI researchers should conduct research on trust in CMC presents itself as a dilemma: Experiments based on a game-theoretic paradigm allow for control, quantitative measurements, and avoid the problems of post-hoc rationalisation. However, they suffer from limited ecological validity and real-world applicability (see Figure 2).
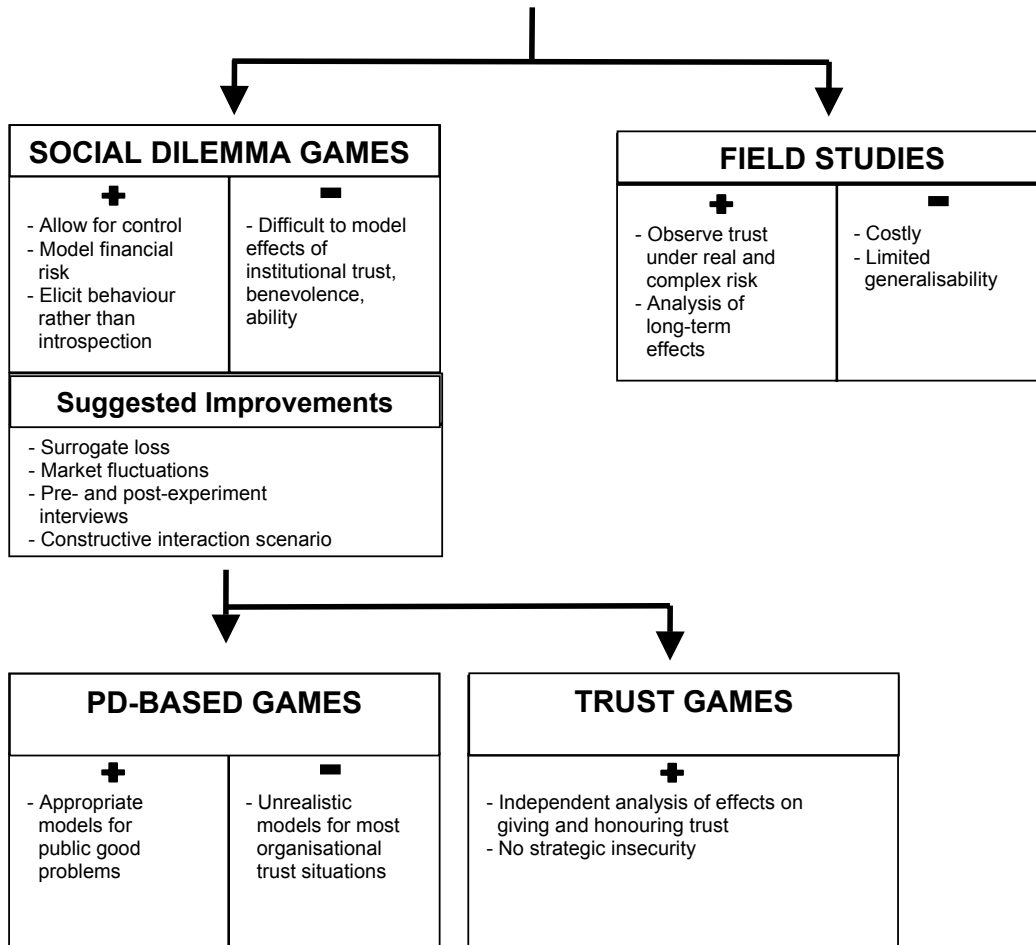
SOCIAL DILEMMA GAMES	
+	**–**
- Allow for control - Model financial risk - Elicit behaviour rather than introspection	- Difficult to model effects of institutional trust, benevolence, ability

Suggested Improvements
- Surrogate loss - Market fluctuations - Pre- and post-experiment interviews - Constructive interaction scenario

FIELD STUDIES	
+	**–**
- Observe trust under real and complex risk - Analysis of long-term effects	- Costly - Limited generalisability

PD-BASED GAMES	
+	**–**
- Appropriate models for public good problems	- Unrealistic models for most organisational trust situations

TRUST GAMES
+
- Independent analysis of effects on giving and honouring trust - No strategic insecurity

Figure 2: Advantages and Disadvantages of the methods discussed and suggested improvements.

PD games are good models for synchronous and symmetric trust situations, such as public good problems. However, in most everyday trust situations, we can identify a trustor, who decides first, and a trustee who then decides to fulfil or defect. Trust Games are better models of such situations. They allow separating the analysis of trust and trustworthy behaviour and they avoid the problems of strategic insecurity.

Social dilemma games in general can be improved by using surrogate risk and through measures that increase the complexity of decisions (e.g. market fluctuations). Furthermore, complementing them with the elicitation of qualitative data, e.g. in constructive interaction scenarios, will help to clarify participant's basis for decisions. Social dilemma games are valuable tools to assess the impact of new technologies in a

relatively simple and controlled way for a large group of situations – mainly those that are based on financial risk and lack of integrity as a source of vulnerability. A prime example is consumer-to-consumer (C2C) e-commerce, offered by auctioning systems such as eBay[6].

However, many of the problems identified cannot be overcome within the game-theoretic paradigm: Experimental studies cannot subject participants to real or unknown risk, and experimental loss is monetary (i.e. mono-dimensional). Furthermore, affective trust bonds that grow over time, institutional trust and specific abilities cannot be modelled in one-off laboratory settings. Thus, results from social dilemma research need to be subjected to field experiments or longitudinal studies that rely on qualitative and ethnographic methods, as they have been employed by Jarvenpaa & Leidner (1998) and Herbsleb, Mockus, Finholt, & Grinter (2000).

So far, CMC-research on trust using PD-based games, confirmed theoretical predictions on distant interaction: Rich media and acquaintance prior to distant interaction help users to develop trust. These conditions now need to be studied with the ecologically more valid methods mentioned above, ideally based on a taxonomy of trust that allows to clearly define an area of applicability. If this step is not undertaken, the field of HCI runs danger of falling into the same trap as social psychologists in the 1960ies, who conducted numerous variations of social dilemma experiments with little relevance for real-world group settings (Poundstone, 1993; Kollock, 1998).

Finally, we have to remember that our efforts to design communication technology for trust are hindered by a paradox. The aim of such technologies is to decreases the cost of distant interaction by replacing face-to-face encounters with mediated ones. However, trust is won through iterations of risk-taking. The decision to meet someone face-to-face can bear a considerable cost and thus risk. The very fact that face-to-face communication is more costly than any technological surrogate gives it an advantage as signifier of one's trustworthiness.

6   Acknowledgements

7   References

ADAMS, A. & SASSE M. A. (2001). Privacy in Multimedia Communications: Protecting Users, Not Just Data. *Proceedings of People and Computers XV*, Lille 2001, p. 49-64.

ADAMS, A. & SASSE, M. A. (1999). Users Are Not The Enemy: Why users compromise security mechanisms and how to take remedial measures. *Communications of the ACM,* 42 (12), pp. 40-46 December 1999.

---

[6] http://www.ebay.com

AXELROD, R. (1980). More Effective Choice in the Prisoner's Dilemma. *Journal of Conflict Resolution,* **24***,* 379-403.

BACHARACH, M. & GAMBETTA, D. (2001). Trust as Type Detection. In C. CASTELFRANCHI & Y. TAN (Eds.), *Trust and Deception in Virtual Societies*. pp. 1-26. Dordrecht: Kluwer.

BERG, J. DICKHAUT, J. & MCCABE, K. (1995). Trust, Reciprocity, and Social History. *Games and Economics Behavior* 10, pp. 122-142.

BOHNET, I. & FREY, B. S. (1998). The sound of silence in prisoner's dilemma and dictator games. *Journal of Personality and Social Psychology*, **38**, 43-57.

BOHNET,I., HUCK,S., TYRAN,J.-R., (forthcoming). Instinct or Incentive to Trust? The Role of Informational Institutions. *Jahrbuch für Neue Politische Ökonomie*. Tübingen: Mohr.

BOS, N., OLSON, J. S., GERGLE, D., OLSON, G. M. & WRIGHT, Z. (2002) Effects of Four Computer-Mediated Communications on Trust Development. *CHI2002 Conference Proceedings*. April 20-25, Minneapolis, US, 135-140.

BOSS, R. W. (1978). Trust and managerial problem solving revisited. *Group and Organization Studies,* **3***,* 331-342.

CLARK, L. & SASSE, M. A. (1997): Conceptual Design Reconsidered - The Case of the Internet Session Directory Tool. In H.Thimbleby, B. O'Conaill & P. Thomas [Eds.]: *People & Computers XII, Proceedings of HCI'97*, Bristol, August 12-15. p.67-84. Springer.

CORRITORE, C. L., KRACHER, B., & WIEDENBECK, S. (2001). Trust in the online environment. In M. J. SMITH, G. SALVENDY, D. HARRIS, & R. J. KOUBEK (Eds.), *Evaluation and Interface Design: Cognitive Engineering, Intelligent Agents and Virtual Reality.* pp. 1548-1552. Mahwah, NJ: Lawrence Erlbaum.

DASGUPTA, P. (1988). Trust as a Commodity. In D. GAMBETTA (Ed). *Trust. Making and Breaking Cooperative Relations.* pp. 49-72. Oxford: Basil Blackwell

DAVIS, J. P., FARNHAM, S., & JENSEN, C. (2002). Decreasing online 'bad' behavior. CHI2002 Extended Abstracts. April 20-25, Minneapolis, MN, US, 718-719.

DEUTSCH, M. (1958). Trust and suspicion. *Journal of Conflict Resolution,* **2***,* 265-279.

DEUTSCH, M. (1960). The effect of motivational orientation upon trust and suspicion. *Human Relations, 13,* 123-140.

DIEKMANN, A. & LINDENBERG, S. (2001). Sociological aspects of cooperation. In N. J. SMELSER & P. B. BALTES (Eds.), *International Encyclopedia of the Social & Behavioral Sciences*. Amsterdam: Elsevier Science.

DÖRING, N. (1998). *Sozialpsychologie des Internet.* Göttingen: Hogrefe.

FLOOD, M. (1952). *Some Experimental Games*. Research Memorandum RM0789. Santa Monica, CA: RAND Corporation.

FROHLICH, N. & OPPENHEIMER, J. (1998). Some consequences of e-mail vs. face-to-face communication in an experiment. *Journal of Conflict Resolution,* **35***,* 389-403.

GALLO, P. S. & MCCLINTOCK C.G. (1965). Cooperative and Competitive Behavior in Mixed-motive Games. *International Journal of Game Theory,* **25***,* 302.

GAUTSCHI, T. (1999). A Hostage Trust Game with Incomplete Information and Fairness Considerations of the Trustee. ISCORE Paper No. 131, Utrecht University.

GIDDENS, A. (1990). *The consequences of modernity.* Stanford: Stanford University Press.

GOFFMAN, E. (1956). The Presentation of Self in Everyday Life. New York: Doubleday.

HANDY, C. (1995). Trust and the Virtual Organization. *Harvard Business Review,* **73***,* 40-50.

HAYASHI, N., OSTROM, E., WALKER, J., & YAMAGISHI, T. (1999). Reciprocity, Trust, and the Sense of Control. *Rationality and Society,* **11***,* 27-46.

HERBLSEB, J.D., MOCKUS, A., FINHOLT, T., & GRINTER, RE. (2000) Distance, dependencies, and delays in a global collaboration. Proceedings of Computer Supported Cooperative Work 2000. New York: ACM

HORN, D. (2001): Seeing is Believing: Video Quality and Lie Detection. PhD Thesis (Psychology), University of Michigan, 2001.

JARVENPAA, S. L. & LEIDNER, D. (1998). Communication and Trust in Global Virtual Teams. *Journal of Computer Mediated Communication, 3*.

JENSEN, C., FARNHAM, S. D., DRUCKER, S. M., & KOLLOCK, P. (2000). The Effect of Communication Modality on Cooperation in Online Environments. In *Proceedings of CHI2000*. pp. 470-477. April 1-6, The Hague, Netherlands.

KAHNEMANN, D., SLOVIC, P., & TVERSKY, A. (1982). *Judgment under uncertainty.* Cambridge: Cambridge University Press.

KEE, H.W. & KNOX, R. E. (1970). *Conceptual and methodological considerations in the study of trust and suspicion*. Journal of Conflict Resolution, 14, pp. 357-366.

KOLLOCK, P. (1998). Social Dilemmas: The Anatomy of Cooperation. *Annual Review of Sociology,* **24***,* 183-214.

KREPS, D. (1990). Corporate Culture and Economic Theory. IN J. E. ALT & K. A. SHEPSLE (Eds.) *Perspectives on Positive Political Economy.* pp. 90-143 Cambridge: Harvard University Press.

LAHNO, B (2002A). Der Begriff des Vertrauens. Paderborn: Mentis

LAHNO, B (2002B). Institutional Trust: A Less Demanding Form of Trust? *Revista Latinoamericana de Estudios Avanzados (RELEA).*

LUHMANN, N. (1979). *Trust and Power.* Chichester: Wiley.

MAYER, R. C., DAVIS, J. H., & SCHOORMAN, F. D. (1995). An Integrative Model of Organizational Trust. *Academy of Management Review, 20,* 709-734.

MCALLISTER, D. J. (1995). Affect- and Cognition-based Trust as Foundations for Interpersonal Cooperation in Organizations. *Academy of Management Journal, 38,* 24-59.

MITRA, A. (2002). Trust, Authenticity and Discursive Power in Cyberspace, *Communications of the ACM*, **45(3)**, 27-29.

MIYAKE, N. (1986). Constructive Interaction and the Iterative Process of Understanding. *Cognitive Science*, 10, 151-77.

MONK, A.F & WATTS, L.A. (2000). Peripheral participation in video-mediated communication. *International Journal of Human-Computer Studies*, 52, 775-960.

NASH, J. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, **36**, 38-49**.**

NORMAN, D. A. (1983). Some Observations on Mental Models. In Gentner, D. A. & Stevens, A. A. (eds.) *Mental models*. Hillsdale, NJ: Erlbaum.

OLSON, J.S., OLSON, G. M. & MEADER, D. K. (1995): What mix of audio and video is useful for small groups doing remote design work? Proceedings of CHI'95, pp 362-368.

POUNDSTONE, W. (1993). *Prisoner's Dilemma.* (2nd ed.) New York: Anchor Books.

POTTER, J. & WETHERELL, M. (1987). *Discourse and social psychology. Beyond attitudes and behaviour.* Sage Publications Ltd., London.

RAUB, W. (1992). Eine Notiz ueber die Stabilisierung von Vertrauen durch eine Mischung von wiederholten Interaktionen und glaubwuerdigen Festlegungen. *Analyse & Kritik,* **14,** 187-194.

RAUB, W. & WEESIE, J. (2000). The Management of Matches: A Research Program on Solidarity in Durable Social Relations. *Netherland's Journal of Social Sciences, 36,* 71-88.

REEVES, B. & NASS, C. (1996). The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. Stanford: CSLI Publications.

RIEGELSBERGER, J. & SASSE, M. A. (2001). Trustbuilders and trustbusters: The role of trust cues in interfaces to e-commerce applications. In B. SCHMID, K. STANOEVSKA-SLABEVA, & V. TSCHAMMER (Eds.), *Towards the E-Society*. Proceedings of the 1st IFIP Conference on E-Commerce, E-Business and E-Governemnt, October 3-5, Zurich, Switzerland. pp. 17-30. Norwell: Kluwer.

RIEGELSBERGER, J., SASSE, M. A., & McCARTHY, J. D. (forthcoming). Shiny Happy People Building Trust? Photos on e-Commerce Websites and Consumer Trust. *Proceedings of CHI2003*, 5-10 April, Ft. Lauderdale, FL.

ROCCO, E. (1998). Trust Breaks Down in Electronic Contexts but Can Be Repaired by Some Initial Face-to-Face Contact. *CHI1998 Conference Proceedings*, pp. 496-502. New York: ACM Press.

ROCCO, E., FINHOLT, T. A., HOFER, E. C., & HERBSLEB, J. D. (2000). *Designing as if trust mattered*. Collaboratory for Research on Electronic Work (CREW) Technical Report. University of Michigan, Ann Arbor.

SALLY, D. (1995). Conversation and Cooperation in Social Dilemmas. A Meta-Analysis of Experiments from 1958 to 1992. *Rationality and Society, 7,* 58-92.

SHACKEL, B. (1991). Usability – context, Framework, Definition, Design and Evaluation. In B. SHACKEL, S. RICHARDSON (eds.). *Humand Factors for Informatics Usability*. Pp 21-37. Cambridge: Cambridge University Press.

SNIJDERS, C. (1993) Trust via Hostage Posting. ISCORE Paper No. 8, Utrecht University.

STRAUSS. A. & CORBIN, J. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. 2nd Ed. Sage, Newbury Park.

TAYLOR, M. (1987). *The Possibility of Cooperation.* Cambridge: Cambridge University Press.

TUCKER, A. (1950). *A Two Person Dilemma*. Lecture at Stanford University. In POUNDSTONE, W. (1993). *Prisoner's Dilemma.* (2nd ed.) New York: Anchor Books.

USLANER, E. M. (2002). *The Moral Foundations of Trust*. Cambridge: Cambridge University Press.

VERTEGAAL, R., SLAGTER, R., VAN DER VEER, G., & NIJHOLT, A. (2001): Eye gaze patterns in conversations: there is more toconversational agents than meets the eyes. Proceedings of CHI 2001, pp. 301-308.

WATSON, A. & SASSE, M. A. (1996): Evaluating Audio and Video Quality in Low-Cost Multimedia Conferencing Systems. Interacting with Computers, Vol. 8 (3), pp. 255-275.

WEIRICH, D. & SASSE, M. A. (2001): Persuasive Password Security. *CHI2001 Extended Abstracts*, March 31- April 6, Seattle, US. pp. 139-140. New York: ACM Press.

WHITTAKER, S. & O'CONNAIL, B. (1997). The Role of Vision in Face-to-Face and Mediated Communication. In K. E. FINN, A. J. SELLEN, & S. B. WILBUR (Eds.), *Video-Mediated Communication*. pp. 23-49. Mahwah, NJ: Lawrence Erlbaum.

WILSON, G. M. & SASSE, M. A. (2000): Investigating the Impact of Audio Degradations on Users: Subjective vs. Objective Assessment Methods. *Proceedings of OZCHI'2000*, December 4-8, Sydney, Australia. pp. 135-142.

YAMAGISHI, T., & YAMAGISHI, M. (1994). Trust and Commitment in the United States and Japan, *Motivation and Emotion,* 18, 129-166.

YATES, J. F., & WATTS, R. A. (1975). Preferences for deferred losses. *Organizational Behavior and Human Performance*, 13, 294-306.

ZAND, D. E. (1972). Trust and Managerial Problem Solving. *Administrative Science Quarterly,* **17***,* 229-239.

ZHENG, J., VEINOTT, E., BOS, N., OLSON, J. S. & OLSON, G. M. (2002). Trust with*out* Touch: Jumpstarting long-distance trust with initial social activities. *CHI2002 Conference Proceedings*. April 20-25, Minneapolis, US, 141-146.