

Constructing Social Systems through Computer-Mediated Communication

Barbara Becker and Gloria Mark

German National Research Center for Information Technology
53754 Sankt Augustin, Germany

E-mail: {Barbara.Becker,Gloria.Mark}@gmd.de

Abstract

The question whether computer-mediated communication can support the formation of genuine social systems is addressed in this paper. Our hypothesis, that technology creates new forms of social systems beyond real-life milieus, includes the idea that the technology itself may influence how social binding emerges within on-line environments. In real-life communities, a precondition for social coherence is the existence of social conventions. By observing interaction in virtual environments, we found the use of a range of social conventions. These results were analyzed to determine how the use and emergence of conventions might be influenced by the technology. One factor contributing to the coherence of on-line social systems, but not the only one, appears to be the degree of social presence mediated by the technology. We suggest that social systems can emerge by computer-mediated communication and are shaped by the media of the specific environment.

Keywords

Collaborative virtual environments, social conventions, virtual communities, social presence, avatars

1. Introduction

Recent studies about social processes in the Internet have begun to concentrate on the question of whether computer mediated communication enables people to build up social relations with other persons despite geographical dispersion [1, 37]. It seems to be still rather unclear if the Internet can support the development of new forms of social structures, i.e. virtual communities, which demonstrate social binding and social coherence comparable to those in real life. Studies that support the assumption that computer-mediated communication generates new forms of social systems [36, 37] are confronted with a more skeptical assessment which raises the discussion of whether the variables used to provide evidence for this are really valid [46]. Critics refer to the absence of commonly-shared life-world perspectives in on-line communities [36], while more optimistic researchers point out that the common background in online-environments is generated by communication [24, 31, 37].

In this paper we present a theoretical framework of how the Internet may function as a means to socially bind people in diverse locations and with divergent life experiences. We discuss the notion of real and virtual communities and list the preconditions that must be fulfilled before a group of actors can be regarded as a community. As a starting point to investigating this notion of virtual community, we set out to observe behaviors in various virtual environments and we chose one precondition that we feel can be captured through empirical observation: social conventions. However, social conventions encompass a wide range, spanning from communication rules which serve to establish a common context for members of a community [43] (discussed in section 2), to interpersonal

behaviors manifest in everyday exchanges which serve to coordinate interaction [13]. The latter are more easily observable, and our empirical findings describe how these interpersonal behaviors are expressed differently depending on the virtual environment. Returning to our earlier point, that Internet technology may influence social binding, this leads us to further examine the role that technology plays in shaping such behavioral conventions. We contrast two alternative hypotheses: 1) that the technology creates a sense of social presence which influences behavior, and 2) that people use the available functionality that requires the least cognitive effort to achieve their goals. Lastly, we discuss how our results are a building block toward the larger notion that communication strategies can be developed through computer-mediated communication, and they can aid people toward developing a feeling of group cohesion and individual belonging in these online-communities.

2. Theoretical framework: Characteristics of social systems

2.1. Technology and fragmented societies

Modern western societies are characterized by a strong tendency towards fragmentation and individualization [15, 22]. Traditional contexts and milieus, like social classes or peer groups, no longer function as a kind of environment where identity development takes place and where people are embedded in solid interaction structures. A common ground, developed by general norms or by a commonly shared life-world, seems no longer to exist. Plurality and the diversity of perspectives are typical characteristics of post traditional societies. Consequently, identity increasingly becomes a product of individual ways of inventing oneself. In addition, social binding emerges in different subgroups and milieus which form a background for these self creation processes [40] and which are often described as incommensurable to each other.

Fragmented and individualized societies are confronted with the problem of how to integrate different perspectives and lifestyles to enable comprehension and dialogue. Several arguments and positions have been developed to answer this open question. On the one hand it is argued e.g. by Habermas [19], that general norms have to exist which form a normative fundament to which every member of a milieu can refer to even if the lifestyle principles of the specific milieus are very different from each other.

On the other hand, it is proposed, e.g. by Lyotard [33], that we have to accept the incommensurability of different milieus without looking for a kind of general focus or viewpoint. Others, in the tradition of Luhman's system theory, point out that we need a kind of general negotiation system, which allows a kind of interaction between the different milieus and social systems by developing strategies and rules which enable them to interact with each other and to find transcontextual viewpoints.

Even if all these theoretical viewpoints are worth discussing in more detail, we would like to present here another idea. We assume that interactive media, like the Internet, can serve as a medium which emphasizes the development of new forms of social binding. According to this assumption, technology is used to establish fragile and fluid social structures beyond the diversity and plurality of milieus where people come from in real life and despite all divergent individual perspectives. Technology forms a kind of communication framework which allows, despite all differences in perspectives and lifestyles of the participants, a kind of communication which can produce weak social binding on a transcontextual level. Following this, we may say that the Internet functions as a medium which allows a kind of social integration, because the commonly used

technology forms the basic framework of communication to which everybody refers to. The handling of many Internet communication technologies is rather transparent and easy to learn, so people from different milieus and with different capacities concerning their cultural capital [4] are able to use it. Of course it is necessary to have the financial capital to access such technologies, which is not available still in most parts of the world. Yet for those who have access to the technology, the Internet can be regarded as a medium which constructs new forms of sociality despite traditional social structures and its boundaries.

2.2. Real and virtual communities

A typical example of such a technologically produced form of sociality are so-called virtual communities. By virtual communities, we refer to interactive environments such as MUDs, MOOs, and 3D graphical systems. Virtual communities may be interpreted as fragile social structures who support, on a global, locally disembedded level, common and transcontextual viewpoints and perspectives [46]. These social structures are weaker and more unstable than traditional forms of communities because the common perspectives are not rooted in a concrete commonly shared life-world. The common viewpoints and binding aspects within these virtual environments must be built up by communication again and again, so they show a high fluidity and fragility. Furthermore, within these virtual social structures, communication has to generate significance and meaning which in locally embedded social structures or in traditional social forms emerges by shared life-world perspectives and inherited perspectives and habits [1, 31].

From this perspective, we looked at virtual communities in the Internet. We have chosen collaborative virtual environments as a field of exploration to find out in which way a kind of common basis is generated within these environments and in how far people refer to the same context of meaning when they are entering the space. We wanted to investigate how communication creates social binding, and if this kind of social binding is comparable to that in real life communities.

Before looking at virtual communities in more detail, a further look on sociological research about characteristic aspects of communities seems to be appropriate. According to a number of sociologists and philosophers [15, 22, 28, 29, 35, 42] social communities are based on some preconditions which have to be fulfilled before we may speak of a community. These are:

- identity persistence of the members,
- commonly shared normative fundament
- existence and stability of social conventions
- a common interest
- a collective rationality
- being rooted in the same geographical/local place
- continuity of the group.

The question arises whether these characteristics can be found in virtual communities. We already mentioned above, that according to our assumption in virtual communities commonly shared viewpoints and meaning have to be created in the process of communication [24], because they have not emerged by the embeddedness in the same life-world, by traditional ways of interacting, by common lifestyle and language, or by inherited incorporated habits [4, 43]. It was our goal to observe virtual communities to

look for evidence of the existence of such characteristics described above, and as a starting point we began by focusing on one such precondition, namely, whether we could discover the use of social conventions. Accordingly, our empirical research focuses on one aspect: it was our intention to explore in which way this common background is created within these on-line environments, how social conventions are generated by communication, and in which way the technology forms and influences these conventions.

2.3. What are social conventions?

Before looking at three different virtual environments from this perspective, we should discuss what is meant by social conventions. Especially in social philosophy, social conventions have been described as normative rules of conduct which are based on implicit ethic imperatives [18, 44]. According to this, social conventions are accepted by group or community members even if they have the opportunity to behave in a different way. Social conventions not only determine how to behave within a group, but furthermore, they define some behavior as incorrect. Following this, they guarantee the stability and consistency of a social system. Normally, a distinction between implicit and explicit social conventions has been made in social philosophical discussions [26, 43]. Some social conventions are articulated by explicit agreements, or even laws, which have been established by institutions or responsible persons. However, more often, social conventions are implicit. They determine the behavior of members of a social system without being codified or formulated. Therefore, we assume that an investigation about the use of such implicit social conventions would give insight into the social practice of a social system, i.e. demonstrating the way how people behave and act [47]. Furthermore, social rules are the underlying preconditions of communication [45, 48], because the way people communicate with each other is embedded in social practice and specific life styles, which are determined by implicit social conventions. According to this, social rules function not only for comprehension but also for coherence within such a system by establishing a common context [44] and a common normative fundament.

Social philosophical research has pointed out that new members of a specific social system have to become aware of these implicit social conventions [43, 45]. By learning and accepting them, they will be integrated in such a group or community. Furthermore, one will only be able to communicate with and understand a partner if one has gained some experience about these implicit conventions. Thus, if we regard collaborative virtual environments as specific forms of social systems, it seems to be a successful research strategy to explore the implicit and explicit social conventions as a first step toward gaining an insight into the particular social practice within such environments.

Other empirical studies have addressed social behaviors in virtual environments, such as the nature of turn-taking and avatar movement [5], dynamics of virtual meetings [6], movement in the virtual world [17, experiences from a mixed-reality environment [3], identity construction [11], cultural formations [37], communication in on-line communities [29], and observations in text-based MUDs [10].

3. Methodological approach and research setting

We employ an approach using ethnomethodology [14], whereby through observation, the social conventions which guide the behavior and attitudes of members of a social system can be identified. In ethnomethodology, social systems are regarded as a net of meaningful behavior, not only governed by formal rules and explicit conventions, but

which are guided more often by implicit conventions which are to some extent open, contingent, and flexible. Through the description of observed single phenomena, empirical events can be explained, rather than attempting to identify global structures or formulating general laws. Therefore, we concentrated our research on obtaining detailed descriptions of conventions in communication to get some insight into the social practices of these environments. We selected a set of social conventions to observe that we feel are important regulators in face-to-face communication, which are described in the next section.

Three different on-line environments were chosen in which to study the existence of social conventions: Active Worlds¹ (**AW**), Onlive Traveler² (**OT**), and LambdaMOO³ (**LM**). All environments are accessible from the Internet. These environments were chosen primarily since they have been in existence for some time and offered different functionality for communication and representation, and thus, we expect, for the expression of social conventions. The main differences are that LM is purely text based for both representation and communication, OT has graphical 3D representations and offers text and audio for communication, and AW has graphical 3D representations and offers only text for communication. The basic functionality available for the representations and communication is described in the Appendix.

Three different researchers spent time observing three different on-line environments. Approximately 59 hours were spent in total observation time: 21 hours in AW, 20 hours in OT, and 18 hours in LM. Each observer was primarily responsible for making observations in one particular environment, but all observers also spent time in each of the other environments to become familiar with them. Although the on-line characters adopted by the researchers were varied somewhat, most of the time the same on-line characters were used during the time spent in the environments. The observation was performed during May-June, and October, 1997 for LM, and September-October, 1997 for OT and AW. The observers took notes and recorded behavioral observations under assigned categories of social behaviors, described in the next section. The observers met periodically and compared observations to make sure that the categories were being coded consistently. On-line recording and logging was not performed due to privacy concerns.

4. Results

We had chosen a set of social convention behaviors to observe that, according to Schefflen [39] serve a regulatory function among actors by initiating, coordinating, and closing interaction. The results reported here are part of a larger study in investigating social behaviors in virtual environments. For a more detailed description see [2].

4.1 Contacting others: greeting and acknowledging

The first convention we address is that of contacting others: greeting and acknowledging. We focused on such a convention since the form of a greeting can influence subsequent interaction. Further, in a virtual environment greeting rituals could be carried out in a number of ways or may not exist at all. In all environments, informal greetings were regularly used to initiate conversation. Yet the form of approaching another and greeting took on different forms.

¹ Copyright © 1997 COF inc.

² Copyright © OnLive! Technologies, 1997.

³ Designed by Pavel Curtis at Rank Xerox Research Centre, 1991.

In OT, greetings are usually directed at individuals, or to a specific group. The greeting is usually audio-based, and the initiator of the greeting usually repositions the avatar to face the other. Greetings are not made when one first enters the world but it is often observed that avatars initially scan the scene (by rotating or moving around). The avatar then navigates to a position close to another before it initiates a greeting. Reciprocity in greetings was also found. If the observer's avatar is not already positioned directly in front of another avatar, the other will turn to face the observer, in the same way that Goffman [16] describes as becoming engaged in talk through face-to-face contact. In fact, sometimes considerable trouble was taken to reposition the avatar. Actors first respond with audio, when the audio is working and quality is good. Once when a person took a long time to respond, he apologized saying he was overwhelmed with text messages. Smiles (shown on the avatar's face) were not observed to have an effect in initiating conversation (nor was it observed to be reciprocated).

In OT, new contacts were made by moving the avatar to face another and addressing the other with audio. When an avatar is spatially very far away, they are generally not approached by other avatars. This was observed with other avatars and tested by the observer who positioned herself far away. The observer received several text messages in greeting, but was never approached by an avatar. The face-to-face positioning during interaction is a convention also found by Bowers et al., [5] in a virtual environment where audio was used.

In AW, greetings are first made as more of a public greeting, to all in the room (only a set of avatars who are close see the greeting. Greetings are usually made by the person at the time the figure joins the location. Only about 30% of the time does an avatar move close to another to face it, when a greeting is made, and as the conversation continues. Private greetings may be made to individuals after, using the avatar name. Reciprocity was also found, but the response to a greeting is not from those avatars in closest proximity but from anyone within this group of 12 avatars, generally two or three others. Gestures for greetings, in the form of an avatar hand-wave, was returned a few times when initiated by the observer, but the observer never received a wave from another as greeting.

Similarly, in LM contacts are first made as more of a public greeting, to all in the room. The whisper command in LM may then be used which allows private communication. Acknowledgment is made also only by a few in the room. Text descriptions of facial expressions and body gestures in LM are sometimes used as greetings (i.e. emote smile). These are often acknowledged by others. For newcomers, a convention is used, following a description in the tutorial, that one announces "Hello, I am new here". People often offer their assistance in response.

4.2 Commitment to a speaking partner

In face-to-face conversation, uninterrupted flow is one type of social rule that is agreed upon between speaking partners and serves to govern conversation [13]. We observed that social behaviors differed in the environments for remaining with, and changing speaking partners. In OT, commitment to a speaking partner was certainly influenced by the face-to-face stance of the avatars. The observer noticed that she herself felt a social obligation to remain for a short time speaking with another, once the face-to-face avatar contact was established. When speaking partners parted, generally a farewell was exchanged between the individuals or other members of the small group.

In AW, the avatars did not change their position very often when new contacts were made. It was possible to change communication partners by simply typing in a new avatar name in the public text window. Speaking partners appeared to change more often in AW than in OT indicating less time was spent with each partner. Thus, using an avatar's name in the greeting signaled that the message was for a specified avatar. It was observed that farewells were said to the public group.

4.3 Group interaction strategies

In face-to-face conversation, spatial positioning indicates who is clustered in a group. A number of social rules exist to govern group interaction: agreements on spatial territory [32], the closeness of members [9], and common group behaviors [34]. We were interested to see what type of behaviors we could observe when actors were conversing in groups in the virtual environments.

In OT, the avatars' graphical positions give information about who they are interacting with. When a group exists, actors generally welcome one into the group by repositioning themselves to form a circle thereby including the new member. One sees by scanning the environment, who is interacting with whom, and the size of the group interaction. It is rude to simply barge into the middle of an existing group. When one approaches a group, the actor generally rotates their avatar around to see whether they are blocking another. Sometimes one will pull their avatar far back to see the complete positioning of the group members, a way to compensate for the lack of peripheral vision of the avatars. New visitors to OT (confirmed by asking them) are often characterized by coming into the middle of the group and not looking around. When the observer or others did this, it provoked annoying reactions.

In AW, since actors reposition themselves less often to face another, or to form groups, group membership is determined by the text flow in the scrolling window, i.e. who is talking with whom. Thus, the visual information becomes less important than the text for this purpose. The observations of Kauppinen et al., [25] confirm these observations, adding that in AW the lack of repositioning the avatars led to confusion. Sometimes the avatars would be layered on top of each other, and with similarities in costumes, identification became difficult. In addition, sometimes the text dialogues, which all appeared in the public window, became too complex to splinter up into different group conversations.

In LM, group membership is often unclear. Only by observing the text dialogues can one ascertain who is talking with whom to get an insight into the interaction structures within the environment. However, if people are conversing with the whisper command, group membership is completely unknown (which can also happen in OT and AW when private messages are sent).

4.4 Signaling privacy

One of the most common ways of signaling private conversations in face-to-face environments is through spatial positioning; speaking partners separate themselves physically from others [20]. Chat rooms on the Internet are based on the model of physical architecture, offering private as well as public rooms. In the environments we looked at, there were also additional methods of indicating privacy: sending private text messages.

Yet we observed that especially in OT, people took advantage of the graphical information in the environment to remain in the same large space and still engage in private conversations. For example, two avatars were once turned completely upside down to signal that they were having a private conversation (with their own common perspective). This was confirmed when the observer (who was right side up) approached them, tried to join in, and was not acknowledged. Joint movement can also indicate privacy, e.g. moving below the floor to a semi-hidden location. Absence of lip movement in avatars facing each other generally means they are having a text conversation, and this is often an indication that it is private, since the observer was generally not acknowledged. (Note: the avatars may also be disconnected from the system, but then the avatars vanish after about a minute). Two avatars conversing far off in the distance from the main arena also signal a private conversation.

In AW, avatar positioning is sometimes also used to indicate a private conversation. This was observed when two avatars were positioned face-to-face very close together. But privacy can also be arranged when the actors would move to another location where others cannot see their text messages. This would be done by moving away, or teleporting to another location. Private telegrams can also be used (text messages), but this function only exists for paying members.

In LM, because no visual information about dialogue situations exists, people can create their own private spaces without being seen by others (using the whisper command). Thus, one is not aware of disturbing intimacy of others.

In all environments, when avatars are engaged in a private conversation, the reaction to any attempt to enter into the conversation is to simply ignore the outside party. Privacy could be signaled by visual means, by positioning the avatars, by changing the communication channel (as in OT), and even by changing language, as observed by Kauppinen et al. [25].

4.5 Interpersonal distance

In the physical world, people maintain a distance from other people during interpersonal communication which serves as a zone of comfort [12]. Evidence that interpersonal distances were perceived was found in both OT and AW. Similar results that confirm these observations were also found in [25] and [23].

In OT, positioning an avatar too close to another provoked annoyed responses from these actors that suggest that they felt their social distance is being violated. This implies that a perception of such an interpersonal space exists. Sometimes avatars in OT moved very far away quickly in the distance as a response or turned to the side. The reactions to closeness could also be due to blocking one's view by the avatar. The observers tested this hypothesis by moving close to others on the side without blocking the view, but the same reactions were observed.

In AW, similar types of reactions were also observed when avatars would come too close to one another. In AW, the text above the avatars overlaps when the avatars are too close (text also appears in the window below). The comments suggest that it is not the text overlap that people are annoyed about, since their comments address that a social distance being violated, e.g.: *You're too close, I can't breathe.*

In LM, interpersonal distance was expressed through text, e.g. emote: comes close, but such commands seldom occurred, and no reactions were observed.

5. How does technology influence social conventions?

The empirical observations reveal a number of social behaviors in these virtual environments that we consider to be conventions in that they fulfill a regulatory function in interaction. Our hypothesis, according to which technology creates new forms of social systems beyond real-life milieus, includes the idea that the technology itself may influence the way how social binding emerges within these on-line environments. We assume, that the specific media and functionality which is available will influence the way that a common background is generated, which social conventions emerge in the communication process, and whether these new forms of social binding are stable or not.

Yet it is not yet clear how the technology might exert an influence. We consider two different explanations that can explain the role of technology in influencing behaviors. On the one hand, the technological environment may be perceived as a window to a shared space, or "portal" as suggested by Kauppinen et al. [25], which connects people to each other. Then, depending on the "clarity" of this portal that the technology affords, people would perform those actions that they would do normally when believing they are in the presence of other people. Another explanation concerns the nature of the technology itself; the handling of the specific media and functionality may lead people to perform certain actions. According to this idea, people choose that functionality that enables efficiency. We begin by discussing the first explanation more in depth.

5.1 Social presence in virtual environments

Although in many ways we can argue that the conventions in the different environments are comparable, the specific behavioral expressions differed. Since the existence of such regulatory behaviors suggests that people are trying to develop on-line communities, it raises the question, why are various conventions used in different on-line environments? In other words, although the Internet offers a common basis for communication, we observe that communicative acts are expressed in different ways.

Our clue to this answer is that all of these environments offer different media and functionality for communication, navigation, and representation. It is our hypothesis that social conventions in such virtual environments are more socially binding if the technology supports a sense of social presence of other actors. This idea refers to social presence theory [41] which states that the nature of the media has an effect on the type of interaction. The stronger the perception of nonmediation in the environment, the stronger is the feeling of *presence* [30]. Social presence is a perception of others that is enabled by a particular technology. Presence thus becomes an interim variable which mediates interaction, and specifically for our study, the expression of conventions. As Short et al. [41] describe, audio-only (and text) media fail to convey a number of visual cues present in face-to-face interaction, such as facial expression, eye gaze, gestures, and proximity. And where important visual cues such as gaze are missing, and which serve as coordination devices for face-to-face partners, we would expect that in such situations interaction would be distorted, compared to face-to-face. The degree of social presence is determined by the conveyance of a number of such non-verbal cues by the media, which influence how present or distant one feels from another person. A high degree of presence suggests the illusion that one is directly interacting with another, and the medium becomes less apparent [30].

Thus, we would expect that the greater the ability to communicate a range of nonverbal cues in a virtual environment, the stronger the sense of social presence that would be created. Of course task is a variable that influences the degree that people rely on nonverbal information; for example, problems of an intellectual nature are generally expected to rely less on nonverbal cues. Yet in the environments that we investigated, the tasks were uniform: socializing and meeting people, which is affected greatly by nonverbal cues.

How social presence might be conveyed in these environments is not so clear-cut. Table I presents a summary of the different media and functionality available in these environments. A more detailed description of the functionality in the environments is presented in the Appendix.

CVE	Communication channels	Representation of actor	Navigation	Representation of environment	Nonverbal Cues
OT	Visual (3D) + audio	Visual avatar	Visual, with mouse and keyboard	3D graphical	Avatars have set of standard expressions; eye-blinks and lip movement
AW	Visual (3D) + text	Visual avatar	Visual, with mouse and keyboard	3D graphical	Avatars have standard gestures; random motions
LM	Text	Text description	Commands with Text	Room metaphor, from Text	Emote command

Table 1. Different media and functionality available in the CVEs observed.

On the one hand, based on media research which shows that visual media facilitate more presence than audio, which has more presence than text media [41], we would expect that OT, which contains visual and audio media, would facilitate more presence than AW, which contains visual and text media. This in turn, would facilitate more presence than LM, which contains only text media. Yet this prediction is made more complex by the fact that in all environments, functionality exists to convey some type of nonverbal expression. We see in Table 1 that in OT, people can activate the avatar to show one of four standard emotions. In AW, the avatars can also be activated to show one of four standard gestures, and in LM, an emote command is designed for expressing emotions.

However, the observers discovered that these "pre-canned" avatar expressions were seldom used; instead people conveyed emotions and expression through the communication media. In OT, emotions were expressed instead via speech, e.g. laughing or with an utterance such as a sigh. In one user's words, "*when you laugh, that says a lot*". In AW, emotions were rather communicated with text: e.g. :o) or *blushing*. The use of emoticons were common, and were also used with LM. In LM, emotions are also expressed both with the "emote" and "say" commands.

It is true that in the graphical environments the avatars show random movements, e.g. blinking their eyes, or folding arms, but it is the observers' agreement that after a short time watching the avatars, these movements did not convey much nonverbal expression.

Thus, according to social presence theory, we would still expect that OT actors would experience the greatest amount of presence due to the graphical information and audio, AW actors a moderate amount due to the graphical information and text, and LM the least amount of presence due to the pure text medium.

5.2. Hypothesis 1: Conventions shaped by a sense of social presence

We re-examine our results of the differences in how on-line conventions are expressed, according to how a social presence hypothesis might explain the results.

Contacting others: greeting and acknowledging

According to social presence theory, actors in OT moved close and faced each other, but not in AW, because the audio channel in OT created a stronger feeling of presence than in AW. Spatial audio forced the actors to come close enough for the audio output to be clear, and the sense that the others were present and "inhabiting" their avatar led people to rearrange their position to face the other.

Commitment to a speaking partner

Conventions differed in the environments for changing speaking partners. According to social presence theory, the face-to-face stance in OT combined with the audio media would lead people to become more engaged with others in conversation in OT, compared to AW. And actors were observed switching conversation partners more often in AW. Just as in a real cocktail party, people may move from one group to another, but social pressures exist for people to spend time with another person in conversation, without leaving too abruptly.

Group interaction strategies

The careful repositioning of OT avatars to make room for a new member in the group's circle can be explained by a feeling of social presence. Along a similar vein, the lack of repositioning in AW when conversation partners changed is consistent with a lower sense of social presence. In fact, the confused layering of avatars that Kauppinen et al.[25] reports supports the idea that users in AW do not behave as though they strongly believe that their avatars are "inhabited".

Signaling privacy

A sense of social presence in OT and AW would have led people to move away from others to engage in a private conversation, since it is impolite to speak privately in front of others. However, due to the nature of our methodology, we did not measure the exchange of private text messages (which could be done, e.g. through logging techniques); thus, we cannot judge the amount of private conversations in the environments. But we can say that private conversations did take place in all environments. Perhaps even a weak sense of presence might trigger the desire to meet with another privately.

Interpersonal distance

Social presence theory would predict that such reactions to violations of personal space would be stronger in OT where presence should be greater. But in fact, such reactions occurred in both OT and AW. But a closer examination shows that in violating

interpersonal distance, conversation exchange (i.e. an audio or text channel) is not involved. Only moving avatars too close to one another results in a violation of the interpersonal space, and this is conveyed through the visual channel. Thus, these actions are not contrary to social presence theory since the act involves only the visual media which is the same in both of the environments.

5.3 Hypothesis 2: Cognitive ease in handling functionality

Although social presence theory accounts for why some conventions are used, it does not fully explain how technology might mediate the formation and use of conventions. We turn now to an alternative explanation which is parsimonious, that concerns the interface design. According to this explanation, the interface design and functionality and media in each of these environments influenced the actors in their behaviors. For example, spatial audio in OT would force an avatar to move close to another during conversation; otherwise, actors could not hear each other, or must send text messages. In AW, conversation is mediated with text and moving close to another avatar that one is communicating with is not necessary. This explanation involves the notion of *cognitive ease*: functionality is used in such a way that it requires the least cognitive effort to reach the goal. This view is based on the model of a user who strives to conserve limited processing resources [7].

Contacting others: greeting and acknowledging

Since the audio in OT is designed for spatial perception, the avatars must move close together to hear each other. If actors want to communicate with text, they may, and still remain spatially distributed, but with our observation methods, we could not determine how many actors were communicating with text. It was the observers' own experience that text was used when the audio quality was poor, and even then, avatars usually faced one another. But a cognitive ease explanation does not address why, when simply moving close activates the spatial audio, actors sometimes went to considerable length to face one another. Cognitive ease would certainly explain in AW why the avatars generally did not face each other to greet and reposition when they continued the conversation. It was simply less effort to write a new name in the text window than move the avatar.

Commitment to a speaking partner

Cognitive ease applies here as it does with greetings. It would predict that in AW it was easy to simply change the avatar name in the public text window, in order to change the conversation partner; it is not necessary to also change the avatar position. In OT, when one wants to use audio, the actor must manipulate the avatar to another location, which requires more effort. Therefore it is easier in OT to stay longer with the same conversation partner, since there is a cost involved in using the functionality to switch partners.

Group interaction strategies

Cognitive ease does not explain why in OT, the actors carefully positioned themselves into group formations. And for the same reason described above in AW, it is less effort to determine group membership by looking at the chat window than by repositioning the avatars to form a configuration that indicates group membership.

Signaling privacy

According to cognitive ease, it is easy to have a private conversation simply by changing the communication media in OT, e.g. from audio to text. An argument against cognitive ease is that it is more effort to signal privacy in OT through graphical means, such as by turning the avatars in a private group upside down, or by moving to a distant location.

Interpersonal distance

Cognitive ease would not explain why reactions to violations of personal space occurred, nor can it explain why a perception of interpersonal distance appears to be transferred from the physical environment to the virtual.

6. Discussion

So far in this paper we have argued that the presence of social conventions supports the notion that virtual environments have emerged as a new form of social system for geographically disperse people. We have discovered that conventions exist in all the environments we observed, but are expressed differently. This led us to further explore different hypotheses for how technology in a virtual environment might mediate the formation and use of conventions: 1) that a particular technology facilitates a sense of presence that others are really in the same shared space, and 2) the behaviors result from handling the available technology to efficiently navigate and communicate.

6.1 The influence of technology on the expression of social conventions

In evaluating the different hypotheses, an overall conclusion is not clear. According to social presence theory, OT should have afforded the greatest sense of presence, leading people to perform social behaviors as if they felt that others were sharing the same space with them. And social presence theory does explain the face-to-face positioning of avatars in OT, as well as accounting for violations of personal space. Especially in virtual environments which offer graphical representations, the nature of these conventions suggests that people seem to identify to a large extent with these representations. They feel more responsible to their conversation partners as evidenced by, for example, explaining why and when they have to leave and reacting sensitively to the violation of personal space. But we also see differences in the two graphical environments. In OT, people appear to behave as though they "inhabit" their avatars, through their care in repositioning themselves and facing each other when speaking. In AW, the avatars seem to function more as a marker, especially for navigation. Also, the expression of a social distance suggests an identification to some extent of the physical body with the graphical representation. It also suggests that the space in the virtual environment is understood and translated as a space similar to that in the physical world, one which contains a particular set of behavioral expectations [21].

Cognitive ease on the other hand makes sense for explaining how conversation partners are changed in AW (i.e. in the text chat window). If changing partners is as easy as typing in a new name, then it is not worth the effort to navigate to a new location to interact with another (as long as one can see them). Considering this result together with OT, we therefore propose the following which takes both hypotheses into consideration: an environment which conveys a high level of social presence will lead people to naturally apply social behaviors that they use in face-to-face interaction. And users will try to use the technology to mimic such behaviors. On the other hand, when this feeling of social presence is low or lacking, then there is less social pressure to follow a face-to-face interaction model so closely. Conventions do exist nevertheless; however, we argue that

their expression arises from the amount of social presence in conjunction with how the functionality and media can be used in the environment.

But none of these explains the most fundamental finding: *that conventions exist at all in online environments*. For this reason, we argue that the existence of conventions supports our hypothesis, that virtual communities have to establish a kind of common background to which people can refer beyond all individualistic or milieu-specific differences. This common normative background is established by communication to overcome the lack of shared life-world perspectives within these environments. Our findings support our assumption according to which social coherence can only be built up in online communities if people can refer to shared beliefs and common interests. As these do not exist by being rooted in the same life-world or by living in geographical or intellectual neighborhoods, communication strategies like social conventions must substitute for this absence. In fact, the use of social conventions are widespread in many Internet environments, such as in newsgroups, one example being the avoidance of capital letters which indicate shouting. In a survey of newsgroup users, most responded positively that they felt a sense of belonging and feeling of closeness to other newsgroup members, which Roberts [38] argues as evidence that a sense of community is felt across many types of newsgroups. The fact that conventions emerge on the Internet when only a text channel is used, shows how strong the urge is for users to establish conventions as a form of regulating and establishing common communicative behaviors.

6.2 Shaping culture through the virtual environment

Just as environmental factors in the physical world shape culture such as, e.g. climate, terrain and natural resources available, we should also expect the technological environment to shape the culture of its inhabitants as well. The design of the virtual environments may contain appropriate metaphors and cues that guide users to act in certain ways. When we consider that often people transfer the conventions that they use in interaction in the physical environment to technology use [8], then the metaphors and cues in the environments could trigger the use of specific conventions. As mentioned earlier, in text-based newsgroups many linguistic conventions have developed, ranging from abbreviations, to determining ways for authenticating user identity and information through writing conventions [11]. Thus, we should also expect other facets of technology to shape culture and influence the formation of conventions as well. A good example of what we have seen in this respect is using the technology to position the avatars in unusual ways, e.g. upside down, to signal privacy. Another example is the availability to switch communication channels to engage in a private conversation.

The expression of emotion was also to some extent shaped by the media available. In OT and AW, users were given a choice for nonverbal expression: audio (OT), text, or changing the avatar expression. In all three environments, emotion was conveyed through the media that provided the most expressiveness. It is more direct and natural to express emotion through speech via the audio channel in OT, than to activate an avatar expression. Further, speech provides a stronger, more individualistic, and more finer-nuanced expression than a "pre-canned" standard avatar look. Similarly, with AW, the text media provides a richer way to express emotion than a standard gesture, even when emoticons and linguistic conventions (e.g. LOL, for laughing out loud) are used. The emote command was used quite often in LM to express nonverbal emotion, and its availability may have encouraged its use.

6.3 Social binding

It is our sense that a feeling of social presence must influence the degree of social binding to the normative behaviors. If we compare these findings with characteristics of news groups, we may say, that in news groups social binding is produced by commonly shared topics and interests while in communities like MUDs and MOOs, this social coherence is more generated by social conventions and social presence.

Undoubtedly certain actors exert a social influence in shaping virtual behaviors, since it is an interactive environment. The presence of "gurus" and expert users in the environments, who often gave helpful advice and tips on using the systems most likely served to influence behaviors by directing people toward certain functionality. And the capability of observing others' actions, which is more apparent in a graphical environment, most certainly also plays a role in spreading codes of social behavior.

Our results demonstrate that virtual worlds can become a kind of specific milieu [40], including characteristic ways of using language, specific interaction modes and particular ways of getting in contact with each other and keeping communication lively. In communication processes, people create a specific meaning within these environments, i.e. they develop a kind of code which is only understandable for frequent participants and which excludes other ones. This seems to be especially true for LM, a virtual environment which has been in existence a longer time than the others. But we may say that in general, social conventions play an important role in developing a specific code of behavior and language which creates social coherence within these online-environments. People who are coming for the first time to these virtual spaces have to become aware of these conventions and follow them to be accepted by the others. In fact, people report being uncomfortable by their lack of knowledge of the conventions. They claimed their messages were not taken into account or that they were treated as outsiders who have to learn how to behave, as one user describes, "*....cause I didn't know the 'in-jokes' and the current word games*".

7. Conclusion

We suggest that our empirical findings can be interpreted as an indication that computer-mediated communication generates and transforms social structures. Even if the social binding within these "virtual" social systems seems to be weaker than in traditional social systems, there exists some sort of group coherence in these communities through establishing shared codified behavior [1, 31]. In addition to other factors, e.g. common themes as in newsgroups, this social binding may also be facilitated by social presence. In these on-line environments, communication seems to be possible even if individuals are not members of the same social milieu and even if they have a different social background. People can, on a very superficial level, begin to communicate with each other without having to refer to the same life-world and shared beliefs. Accordingly, technology and how it is used form a new context which is accessible for people from very different milieus. It enables them to understand each other in spite of these differences. By this, technology may have an integrative social effect, and we propose that it might even counteract the tendencies of fragmentation and individualization in modern western societies.

However, we have to concede that our findings are only a first step in finding evidence for the existence of social structures in the internet. If we look back at what we described as typical characteristics and preconditions of social communities, it is clear that the social conventions we found cannot be regarded as satisfactory proof for virtual sociality. Further studies have to be done and we hope that our initial attempt will provoke further

research in this direction considering our theoretical framework. So, even if it is still unclear whether computer-mediated communication may function as an integrating parenthesis by clustering different social perspectives, we propose that our first findings support the position of sociologists like Knorr-Cetina [27] who states that technology is not only born of social systems but also serves to create and transform social structures.

Acknowledgments

We wish to thank Phillip Jeffrey for his valuable help in making this study possible, and Margarete Boos for her comments on the paper.

8. References

1. Baym, N.K., The Emergence of Community in Computer-Mediated Communication, in: Jones, S. (ed) , CyberSociety. Thousands Oaks, London, New Dehli 1995
2. Becker, B. and Mark, G. (1998). Social conventions in collaborative virtual environments. *Proceedings of Collaborative Virtual Environments 1998 (CVE'98)*, Manchester, June 17-19, 1998.
3. Benford, S., Greenhalgh, C., Snowden, D., and Bullock, A. Staging a public poetry performance in a collaborative virtual environment. Proc. Of the 5th European Conference on CSCW, Dordrecht: Kluwer Academic Publishers, 1995, pp. 125-140.
4. Bourdieu, P., Die feinen Unterschiede, Frankfurt 1987
5. Bowers, J. Pycock, J., and O'Brien, J. Talk and embodiment in collaborative virtual environments. Proc. of CHI'96, ACM Press, 1996.
6. Bowers, J., O'Brien, J., and Pycock, J. Practically accomplishing immersion: cooperation in and for virtual environments, Proc. CSCW'96, ACM Press, 1996, pp. 380-389.
7. Card, S. K., Moran, T. P., and Newell, A. , The Psychology of Human-Computer Interaction. Hillsdale, NJ: Lawrence Erlbaum Associates, 1983.
8. Carroll, J.M. and Thomas, J.C. (1982). Metaphors and the Cognitive Representation of Computing Systems. *IEEE Transactions On Systems, Man, And Cybernetics*, Vol. SMC-12, No. 2.
9. Cheyne, J. A., and Efran, M. G. (1972). The effect of spatial and interpersonal variables on the invasion of group controlled territories. *Sociometry*, 35, 477-489.
10. Curtis, P. Mudding: Social Phenomena in text-based virtual realities. In M. Stefik, Internet Dreams: Archetypes, Myths, and Metaphors, Cambridge: MIT Press, 1996, pp. 265-292.
11. Donath, J., Identity and deception in the virtual community, in: Kollock, P. and Smith, M. (eds). Communities in Cyberspace, Berkeley: University of California Press 1998.
12. Dosey, M. A. and Meisels, M., Personal space and self-protection. *Journal of Personality and Social Psychology*, 11, 93-97, 1969.
13. Duncan, S. and Fiske, D. W. 1977. *Face-to-Face Interaction*. Hillsdale, N.J.: Erlbaum.
14. Geertz, C., *Dichte Beschreibung. Beiträge zum Verstehen kultureller Systeme*, Frankfurt: Suhrkamp Publishers 1983.
15. Giddens, A., The Consequences of Modernity, Cambridge: MIT-Press, 1990.
16. Goffman, E. , Forms of Talk. Oxford: Blackwell, 1981

17. Greenhalgh, C., Analysing movement and world transitions in virtual reality teleconferencing. Proc. of the 5th European Conference on CSCW, Dordrecht: Kluwer Academic Publishers, 1995, pp. 313-328.
18. Habermas, J., Theorie des Kommunikativen Handelns, vol. 2, Frankfurt: Suhrkamp, 1981.
19. Habermas, J., Faktizität und Geltung, Frankfurt, 1992.
20. Hall, E. T. (1966). *The Hidden Dimension*. New York: Anchor Books.
21. Harrison, S. and Dourish, P., Re-Place-ing Space: The roles of place and space in collaborative systems. In M. Ackerman, (ed). Proc. of the ACM 1996 CSCW Conf., Nov. 16-20, 1996, Boston, MA.
22. Honneth, A.(ed.), Kommunitarismus, Frankfurt 1994
23. Jeffrey, P. and Mark, G., Social navigation and personal space: An empirical study in virtual environments. *Proceedings of International Workshop on Personalised and Social Navigation in Information Space (IFIP WG13.2)*, Stockholm, March 16-17. (SICS Technical Report T98:02, 1998), Stockholm: Swedish Institute of Computer Science (SICS), pp. 24-38, 1998
24. Jones, S. (ed) , CyberSociety. Thousands Oaks, London, New Dehli, 1995
25. Kauppinen K., Kivimaki A., Era T. and Robinson M. Producing Identity in Collaborative Virtual Environments in VRST'98: Symposium on Virtual Reality Software and Technology. 1998 (forthcoming). Taipei, Taiwan: ACM.
26. Kemmerling, A., "Regeln", in: Speck, J. (Ed.) Handbuch wissenschaftstheoretischer Begriffe, Göttingen: UTB, 1980
27. Knorr Cetina, K., Sozialität mit Objekten, in: Rammert, W. (Hrsg), Technik und Sozialtheorie, Frankfurt 1998
28. Kollock, P., /Smith, M. (eds.), Communities in Cyberspace, Berkeley, 1998
29. Kollock, P., Design principles for online communities, In The Internet and Society: Harvard Conference Proceedings, Cambridge, MA: O'Reilly & Associates, 1996.
30. Lombard, M. and Ditton, T., At the heart of it all: The concept of presence. *Journal of Computer Mediated Communication* 3 (2), September 1997.
31. Lyons, D., Cyberspace-Sozialität: Kontroversen über computervermittelte Beziehungen, in: Vattimo, G./Welsch, W. (Hrsg), Medien - Welten Wirklichkeiten, München 1997.
32. Lyman, S. M., and Scott, M. B. (1967). Territoriality A neglected sociological dimension. *Social Problems*, 15, 236-249.
33. Lyotard, J.F., Das postmoderne Wissen, Wien 1986
34. Moreland, R.L. and J.M. Levine. Newcomers and Oldtimers in Small Groups. In *Psychology of Group Influence*, P. Paulus (eds.), Lawrence Erlbaum, Hillsdale, NJ, 1989, 143-186.
35. Opielka, M., Gemeinschaft und Gesellschaft, unveröffentl. Dissertation, Bonn 1996
36. Poster, M., The Second Media Age, Cambridge 1995
37. Reid, E., Communication and community on Internet Relay Chat: constructing communities, in P. Ludlow (ed.), High Noon on the Electronic Frontier, Cambridge: MIT Press, 1996.
38. Roberts, T.L., Are News-Groups virtual Communities?, *Proceedings of CHI 1998*, Los Angeles, April, p. 360 - 367, 1998.
39. Schefflen, A. E., Models and epistemologies in the study of interaction. In A. Kendon, R. M. Harris, and M. R. Key (Eds.) *Organization of Behavior in Face-to-Face Interaction*. The Hague: Mouton Publishers, pp. 63-91.
40. Schulze, G., Die Erlebnisgesellschaft, Frankfurt, 1992.

41. Short, J., Williams, E., and Christie, B., *The Social Psychology of Telecommunications*. London: Wiley, 1976
- 42.. Taylor, Ch., "Aneinander vorbei: Die Debatte zwischen Liberalismus und Kommunitarismus, in: Honneth, A.(ed.), *Kommunitarismus*, Frankfurt 1994
- 43.. Waldenfels, B, *Der Stachel des Fremden*, Frankfurt: Suhrkamp, 1994.
44. Waldenfels, B., *In den Netzen der Lebenswelt*, Frankfurt: Suhrkamp, 1985.
45. Wellmer, A., *Zur Dialektik von Moderne und Postmoderne*, Frankfurt: Suhrkamp, 1984.
46. Wellman, B., Gulia, M., *Net Surfers don't ride Alone: Virtual Communities as Communities*, in: Kollock, P., Smith, M., (eds.), *Communities in Cyberspace*, Berkeley, University of California Press, 1998
47. Winch, P., *Die Idee der Sozialwissenschaft und ihr Verhältnis zur Philosophie*, Frankfurt 1966.
48. Wittgenstein, L, *Philosophische Untersuchungen*, Frankfurt: Suhrkamp 1960.

Appendix: The media and functionality in the CVEs

AW: full-bodied avatars can walk and exhibit movements of waving, jumping, and dancing, activated by mouseclicks. Avatars can move in six dimensions by using the arrow keys. Communication between people is text-based by typing on the keyboard. All public messages appear in a scrollable window and also above the avatar head with the avatar name for 30 seconds or until the next typed message appears.

OT: the avatars are heads, and have four different emotions that one can activate by a mouseclick: happy, sad, surprise, and anger. The avatars exhibit what appear to be random blinks. Movement (three dimensions plus rotation in four directions—left, right, forward, and backward) is made by using arrow keys. Communication is audio (outgoing audio is activated by pressing down the control key and speaking into a microphone) or text-based (pulling down a menu, selecting an avatar, and typing into a message which appears on the screen). The text is limited to two lines.

LM: all representation of users and communication is text-based. Different commands are used for communication (e.g. say, whisper, emote), manipulation (e.g. get/take, move), information (e.g. look, who, etc.), and creation (e.g. dig, create), as well as others.