Exploring Communicative Modes for a Virtual Environment: A Social Psychological Analysis

Eva-Lotta Sallnäs
The Royal Institute of Technology,
IPLAB/NADA/KTH
SE-100 44 Stockholm, Sweden
+46 (0) 8 790 66 26
evalotta@nada.kth.se

Anders Hedman
The Royal Institute of Technology,
CID/NADA/KTH
SE-100 44 Stockholm, Sweden
+(0) 8 790 92 83
ahedman@nada.kth.se

ABSTRACT
This paper reports from an experimental study of cooperative work in a collaborative virtual environment (CVE) with sixty test subjects. A method from social psychology, interaction process analysis (IPA), was used for analyzing group dynamics within the CVE in relation to three communicative modes: text-chat, voice, and video conferencing. The study investigated how the communicative mode influenced the quality of the social interaction and how cooperative work in a CVE could be analyzed by using categories of social interaction. It was found that the frequencies of utterances in the voice and video conferencing conditions were significantly larger than in the text-chat condition. The dialogue of the text-chat condition was found to be more task oriented than the dialogues of the voice and video conferencing condition. The social psychological method used was not developed for use within CVE:s, but it was found to work well since the scope of the IPA method is broad and applicable to virtually any small group.

Keywords
Collaborative virtual environment, communicative modes, interaction process analysis, experiment

INTRODUCTION
This paper reports from an experimental study that was conducted in a collaborative desktop virtual environment (CVE). The study aimed at gaining a better understanding of the social dynamics and experiences of people collaborating in such environments as they use different communicative modes. By communicative mode is here meant the means by which people communicate with each other in a CVE. Three different such modes were deployed: text-chat, voice, and video conferencing. In the study presented here the aim was to explore, if the mode of the communication channel influenced the test subjects' task and socioemotional orientation when cooperating in a CVE.

Research on electronic communication has shown how media supporting different modalities vary in capacity for carrying information [20, 7, 18]. Communicating through audio is important when collaborating at a distance and improves both task performance and perceived affordances in comparison to text-chat [15, 20, 21]. In one study [4] various modes of interaction were examined ranging over audio, handwriting, typewriting, video and face-to-face. Results showed that people spoke more than they wrote, and people performed best in terms of time to complete tasks when audio was provided compared to text.

Research has not found such uniform results regarding benefits from using video conferencing. It is usually shown that video does not add significant advantages to task performance compared to audio [22, 1, 17]. Jensen et al [12] examined the nature of cooperation between two players of an electronic text-based game under four different conditions: no communication channel, text-chat, text-to synthesized speech, and audio. They found that immediate forms of communication e.g., audio are important for cooperation and trust in online environments. Jensen points to the need for conducting studies in richer environments with e.g., 3d avatars and researching other aspects of social interaction than cooperation and trust.

BACKGROUND
Social psychology has had considerable impact on research in the field of computer mediated communication–CMC [20, 19, 13, 21]. This paper represents an attempt to explore a method from social psychology as an approach to understanding the social dynamics of CVE:s. It relies to an important extent on Bales Interaction Process Analysis (IPA) taxonomy of social interaction. The IPA taxonomy was not developed with CVE:s in mind, but for physically co-located teams. The findings presented in this paper show that the taxonomy also is useful for analysing social interaction in CVE:s.

Interaction Process Analysis emerged as a way to study group interaction by mapping acts that occur in co-located social interaction [2] and has also been applied in CMC [3, 10, 11]. The IPA evolved through several studies of communication acts in face to face meetings. When interacting, a group is seen as taking part in a problem solving activity consisting of a sequence of acts. These acts give the participant a sense of being part of a joint or shared process. Bales argue that in order to maintain the shared process, people engage in two general categories of communication: task-oriented and socio-emotionally oriented communication. These general categories are also part of a more detailed taxonomy that Bales developed for analyzing verbal behavior.

Experimental studies of small physically co-located groups showed that only a dozen categories were relevant to code in collaborative situations [2]. These categories are meant to be collectively inclusive and mutually exclusive, i.e., every observed communication act can be classified as belonging to exactly one category. The unit to be coded is the smallest segment of verbal or non-verbal behavior to which the observer can assign a classification. Thus complex sentences usually involve multiple code units. Verbal acts within a small-group setting are analyzed into the following low-level categories: give information, give suggestion, give opinion, ask for information, ask for suggestion, ask for opinion, show solidarity, show tension release, show agreement, show disagreement, show tension, show antagonism. From these low-level categories, Bales abstracts four mid-level categories: problem solving attempts, questions, positive reactions, and negative reactions. Finally from these mid-level categories two high-level classes are abstracted: task-oriented communication and socio-emotional communication.

The IPA taxonomy makes it possible to compare groups with each other or a group with itself at different times, i.e., before and after the introduction of some experimental variable. The taxonomy was designed with the aim of capturing aspects of interaction so general that they will be manifest in the behavior of virtually any small group. This latter fact makes it plausible that the taxonomy should also be applicable to group interaction in CVE:s and not only to group interaction in physically co-located groups. Moreover, since the IPA taxonomy is so general with respect to physically co-located groups, it can serve as a real-world yard stick against which to examine small group behavior in CVE:s.

METHOD

Experimental design

A between-groups design was used and the independent variable was a CVE with three different communicative modes:

- text-chat
- voice
- video conferencing

The text-chat was a feature of the CVE-software and it appeared as a window beneath and adjacent to the 3D environment. Since the CVE-software had neither voice communication nor video conferencing as built in features, these latter features had to be arranged externally. In the voice condition, participants communicated through telephones with headsets. In the video conferencing condition, two 21-inch television monitors were used in addition to the telephones with headsets.

Participants

Sixty people participated in the experiment. The participants were students from the Stockholm University and the Royal Institute of Technology along with a few administrative staff members. They were assigned to pairs with a woman and a man in each pair and they each performed a simple decision making task together. Pair members did not know each other prior to the experiment.

The collaborative virtual environment

The CVE was constructed in the ActiveWorlds CVE system (www.activeworlds.com) and had the appearance of a simple exhibition with information points (Figure 1). Access to the CVE was restricted to the participants of the study. Two networked portable PC:s were prepared with clients of the ActiveWorlds CVE system.

These information points had posters with pictures of cars and QuickTime movie clips with model-specific information. Humanlike avatars represented the subjects. Female participants had a traditional female avatar with the following characteristics: brown hair in a ponytail, purple trousers and a T-shirt. The male participants had a traditional male avatar with these attributes: black short hair, white T-shirt, and blue trousers.

Figure 1. The exhibition with information points in the ActiveWorlds environment.

Task

The decision-making task was presented to the pairs of participants as a written scenario:

You have participated in a competition in which you and another participant performed equally well. You therefore have to share the first price: a Volvo car of your choice with

insurance and gas for one year. You will use the car every second month. You will not be able to sell the car. The organizers of the competition now want you to go in to a virtual exhibition and choose the car that you are going to share. You should then decide together which one of you is going to have the car the first month.

The participants were encouraged to discuss this scenario with each other and to interpret it together.

Independent variables

One independent variable was varied: the mode of communication involved in solving the task.

In the video conferencing condition, participants had a TV monitor in front of them and wore headsets enabling them to converse via a telephone connection. The video cameras used to document the interaction served dual purposes and were simultaneously used for the videoconference. The angle of the monitors and the video cameras was about 30 degrees to the participants left side. The video cameras were positioned under the monitor and tilted slightly upwards. This latter arrangement served to minimize the parallax problem, as compared to placing the cameras on top of the monitors. Each participant had the notebook PC with the CVE directly in front of them.

In the voice condition the participants communicated via the telephone connection using headsets and had no video connection.

In the text condition the participants communicated only via a text chat available in the CVE.

Dependent variables

Twenty-seven dialogues out of thirty were transcribed from the video recordings. These dialogues were coded into the twelve categories by the two experimenters independently (Table 1). Detailed descriptions on how to recognize and distinguish the different categories are shown in the listing below. As mentioned earlier, the twelve categories can be classed into four types of acts (A-D). Finally A and B can be classed together as, task-oriented communication, and C and D can be classed together as, socio-emotional communication.

Table 1. Twelve categories are listed with descriptions on how to recognize and distinguish them. These twelve categories can be classed into four acts (A-D). Finally A and B can be classed as task-oriented and C and D can be classed as socio-emotionally oriented communication.

Task-oriented communication:

A. Problem-solving attempts

1. Gives information (GI) = orientation, repeats, clarifies, confirms.

2. Gives suggestion (GS) = direction, implying autonomy for other.

3. Gives opinion (GO) = evaluation, analysis, expresses feeling, wish.

B. Questions

4. Asks for information (AI) = orientation, repetition, confirmation.

5. Asks for suggestion (AS) = direction, possible ways of action.

6. Asks for opinion (AO) = evaluation, analysis, expression of feeling.

Socioemotional communication:

C. Positive reactions

7. Shows solidarity (SS) = raises other's status, gives help, reward.

8. Shows tension release (SR) = jokes, laughs, shows satisfaction.

9. Shows agreement (SA) = shows passive acceptance, understands, concurs, complies.

D. Negative reactions

10. Shows disagreement (SD) = shows passive rejection, formality, withholds help.

11. Shows tension (ST) = asks for help, withdraws out of field

12. Shows antagonism (San) = deflates other's status, defends or asserts self.

Documentation

For all three conditions trials were video recorded using a video mixer in such a way that the participants´ faces, torsos and the interaction in the CVE were captured synchronously on the same tape. The navigation of the participants in the CVE was recorded from the perspective of one subject. Audio was captured using one microphone per participant and the text from the text-chat communication was saved for analysis.

Analysis

The classifications of the coders were compared to ensure that the coding was reliable. The code frequencies of the twelve categories respectively and from the two coders were analyzed with one-way ANOVAs (analysis of variance).

If the data from both coders showed the same results regarding significant differences between the three conditions in the twelve categories then the inter-code reliability was regarded as satisfactory. This analysis showed that both coders had equal results in all twelve categories in all three conditions with one exception. The coders did not have the same results in the category give suggestion (GS), when the text and the video conferencing condition were compared.

RESULTS

A comparison of the frequencies of codes in the twelve categories using one-way ANOVAs in the three conditions show a number of significant differences between the text condition and the voice and video conference condition respectively. However, results reveal no significant differences between the voice and video conferencing conditions with respect to the amount of communication between participants. The mean numbers of code units was almost the same in all categories of both the voice and video condition (Figure 2, Table 2). When performing the task, people were markedly more verbally engaged in the voice and in the video conference conditions in comparison to the text condition. Subjects did not seem as engaged in the discussions in the text condition and those dialogues exhibited numerous misunderstandings.

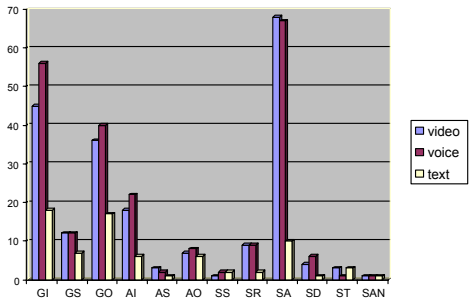

Figure 2. Mean values of frequencies of codes in the twelve categories in the video conference, voice and text condition.

Basic communication categories

One set of categories (GI, GS, GO, AI, AS, AO) is more concerned with task performance. Participants gave information (GI) $F=11$ $p<0.01$, gave opinions (GO) $F=10$ $p<0.01$ and asked for information (AI) $F=24$ $p<0.01$ to a significantly lesser extent in the text condition than in the voice condition. Participants also gave information (GI) $F=8$ $p<0.05$, gave opinions (GO) $F=6$ $p<0.05$ and asked for information (AI) $F=14$ $p<0.01$ to a significantly lesser extent in the text condition than in the video conferencing condition. Participants gave significantly more suggestions (GS) $F=7$ $p<0.05$ in the voice than in the text condition. Note however, that the inter-code reliability was not satisfactory in this single case. Mean values of frequencies of codes in all the twelve categories in the video conference, voice and text condition is shown in table 2.

Table 2. Mean values of frequencies of codes in the twelve categories in the video conference, voice and text condition.

	Video	Voice	Text
GI	M=45	M= 56	M=18
GS	M=12	M= 12	M=7
GO	M=36	M= 40	M=17
AI	M=18	M= 22	M=6
AS	M=3	M= 2	M=1
AO	M=7	M= 8	M=6
SS	M=1	M= 2	M=2
SR	M=9	M= 9	M=2
SA	M=68	M= 67	M=10
SD	M=4	M= 6	M=1
ST	M=3	M= 1	M=3
SAN	M=1	M= 1	M=1

Much of the communication in the voice and video conference conditions concern personal facts such as how large subjects families are, or experiences relevant to the information displayed. Subjects in the voice and video conference conditions also explored the CVE more completely and consequently had more to discuss. Therefore, subjects gave more information (GO) and asked for more information (AI) in the voice and video conference conditions than in the text condition.

Subjects also gave more opinions (GO) in the voice and video conference conditions as they discussed and negotiated more extensively than in the text condition. They often discussed and analyzed the information carefully while at the same time communicating their personal preferences. In the text condition negotiations were more crude.

Another set of categories (SS, SR, SA, SD, ST, San) is concerned with socio-emotional issues. Participants showed tension release (SR) $F=34$ $p<0.01$, showed agreement (SA) $F=34$ $p<0.01$ and showed disagreement (SD) $F=11$ $p<0.01$ to a significantly lesser extent in the text condition as compared to the voice condition. Participants also showed tension release (SR) $F=7$ $p<0.05$, showed agreement (SA) $F=15$ $p<0.01$ and showed disagreement (SD) $F=9$ $p<0.01$ to a significantly lesser extent in the text condition as compared to the video conference condition. Mean values of frequencies of codes in all the twelve categories in the video conference, voice and text condition is shown in table 2.

Subjects showed more agreement (SA) in the video conference and voice conditions that in the text condition. This indicates that it was easier to give feedback in the voice and video conference condition than in the text condition. The difference between conditions with respect to the category show disagreement (SD), reflects that subjects negotiated more in the video conference and voice conditions. This category reveals that it was easier to give feedback about misunderstandings and thus to maintain a common ground in the voice and video conference [6].

Subjects showed tension releasing (SR) behavior more in the voice and video conference conditions than in the text condition. They were joking more and often communicated satisfaction with individual or cooperative behavior. There were examples of tension releasing behavior in the text condition also. Smilies (emotionally expressive symbols created with standard keyboard characters) were sometimes used to communicate jokes or feelings of satisfaction.

Communication acts

In the next step the dialogue was classed into four categories of acts: problem solving, questions, positive reactions and negative reactions in the three conditions and analyzed in a relative manner (Table 3). This relative way of comparing the voice, video conference and text conditions revealed that the main focus in the dialogue was, in all conditions, on problem solving, but most evidently so in the text condition. Moreover, in the text condition, more questions were asked than in the voice and the video conferencing conditions. In the video and voice condition more emphasis was put on conveying positive reactions than in the text condition. The results also show that little verbal communication was about showing negative reactions in any of the conditions.

Table 3. Percentage of acts of problem solving, questions, positive reactions and negative reactions in the video, voice and text-chat conditions.

	Video	Voice	Text
Problem solving	46%	48%	59%
Questions	14%	14%	18%
Positive reaction	38%	35%	20%
Negative reactions	2%	3%	3%

Taskoriented vs. socioemotional communication

The findings were analyzed one step further and the categories were classed together as, task-oriented communication or socioemotional communication (Table 4).

Table 4. Percent of task-oriented and socio-emotionally oriented communication in the video conferencing , voice and text conditions.

	Video	Voice	Text
Task-oriented	60%	62%	77%
Socio-emotional	40%	38%	23%

Communication in all conditions was more concerned with the task than with socio-emotional aspects, most markedly in the text-chat condition. The analysis also reveals that the communication was more socio-emotionally oriented in the video conferencing and voice conditions than in the text condition.

In this phase the analysis of the results shows a tendency that the dialogue in the video condition is more socio-emotionally oriented than the dialogue in the voice condition. The voice condition on the other hand is slightly more task oriented than the video condition.

DISCUSSION

This study has shown that IPA can be successfully applied to the analysis of small group dynamics in a CVE. The categorization was experienced as intuitive and sound by the experimenters and yielded interesting results.

The results presented here are in support of the claim made by Jensen [12] that the richness of the communication channel is an important determinant for the quality of online cooperative work. Communication in the text condition was terse and did not allow the subjects to engage themselves in the more varied verbal behavior evident in the other two conditions. The cost of communicating by using text-chat appeared too great to allow for elaborate discussions or substantial socializing behavior.

The value of voice communication was almost on par with video conferencing. This finding is also supported by earlier work [5, 8, 16]. In comparison to voice, video conferencing was not found to enhance the quality of cooperative work or increase the amount of socializing to a large extent. However, the video channel was used by the subjects in certain situations. The video channel was used when: there were long pauses in the dialogue and both subjects investigated information, when navigation problems occurred, during greetings and during discussions prior to important decision making. Results from the high-level analysis of the categorization into two classes, task- and socio-emotionally oriented communication, do show a tendency that it do matters which of these modes that are used for communication in a CVE.

The reason why more than half of the dialogue in the text condition was focused on problem solving might be that, because communicating in this mode was harder, most of the effort was put into actually solving the task.

The deployment of Bales taxonomy for classifying small group behavior proved to be a viable approach in analyzing the interaction within the CVE under the three conditions. However, as Bales [2] points out the IPA taxonomy was designed with the aim of capturing aspects of interaction so general that they will be manifest in the behavior of virtually any small group.

Issues regarding the specific content and context in a certain case of social interaction, as for example in this specific exhibition in a CVE is not addressed by IPA. From the perspective of Bales [2] a content analysis would be important to conduct in order to obtain complementary information about the specifics of a certain case. The next step would therefore be do conduct a content analysis of the transcripts and video recordings from the study presented here.

ACKNOWLEDGMENTS

Advice and valuable comments from Kerstin Severinson-Eklundh are gratefully acknowledged.

REFERENCES

1. Anderson, A., Newlands, A., Mullen, J., Flemming, A. M., Doherty-Sneddon, G. and Van der Velden, J. (1996). Impact of Video-Mediated Communication on Simulated Service Encounters. *Interacting with Computers*. 8: 193-206.
2. Bales, R. F. (1976). *Interaction process analysis: a method for the study of small groups*. Chicago : Univ. of Chicago Press.
3. Carletta, J., Anderson, A. H., & McEwan, R. (2000). The effects of multimedia communication technology on non-collocated teams: a case study. *Ergonomics*, **43**(8), 1237-1251.
4. Chapanis, A. (1975). Interactive Human Communication. *Scientific American*. 232: 36-42.
5. Chapanis, A., Ochsman, R. B., Parrish, R. N. & Weeks, G.D. (1972) Studies in interactive communication: The behavior of teams during co-operative problem solving. *Human Factors*, 14: 487-509.
6. Clark, H. H., Brennan, S. E., (1991) Grounding in communication. In Perspectives on Socially Shared Cognition. Resnick, L. B., Levine, J. M., and Teasley, S. D. Eds. *American Psychological Association, Washington*. 127-149.
7. Daft, R. L. & Lengel, R. (1986). Organizational Information requirements, media richness, and structural design. *Management Science*. 32: 554-571.
8. Egido, C. (1990) Teleconferencing as a technology to support cooperative work. In Intellectual Teamwork. J. Galegher, R. Kraut, C. Egido, Eds. *Lawrence Erlbaum, Hillsdale, NJ*. 351-371.
9. Fraser, M., Benford, F., Hindmarsh, J., & Heath, C. (1999). Supporting Awareness and Interaction through Collaborative Virtual Interfaces In *Proceedings of UIST '99*, Asheville, NC, 1999. 27-36.
10. Hiltz, S.R., Johnson, K., & Agle, G. (1978). Replicating Bales' problem solving experiments on a computerized conference: *A pilot study (Research Report No. 8)*. Newark: New Jersey Institute of Technology, Computerized Conferencing and Communications Center.
11. Hiltz, S.R., Johnson, K. & Rabke, A.M. (1980). The process of communication in face to face vs. computerized conferences: A controlled experiment using Bales interaction process analysis. In N. K. Sondheimer (Ed.), *The 18th Annual Meeting of the Association for Computational Linguistics and Parasession on Topics in Interactive Discourse: Proceedings of the Conference.* Philadelphia: University of Pennsylvania.
12. Jensen, C., Farnham, S. D., Drucker, S. M. & Kollock, P. (2000). The Effect of Communication Modality on Cooperation in Online Environments. In *Proceedings of the ACM Conference on Computer-Human Interaction (CHI '00),* Amsterdam. 470-477.
13. Kraut, R. E., Fish, R. S., Root, R.W., & Chalfonte, B.L. (1990). Informal communication in organizations: Form, function, and technology. In S. Oskamp & S. Spacapan (Eds)., *Human Reactions to Technology: The Claremont Symposium on Applied Social Psychology. Beverly Hills, CA: Sage Publications*. 145-199.
14. Matsuura, N., Fujino, G., Okada, K. and Matsushita, Y. (1993). An Approach to Encounters and Interaction in a Virtual Environment. *Proceedings of ACM Computer Science Conference*. Association for Computing Machinery. 298-303.
15. Ochsman, R.B. and Chapanis, A. (1974). The Effects of 10 Communication Modes on the Behavior of Teams During Co-operative Problem-solving, International *Journal of Man-Machine Studies*. 6: 579-619.
16. O'Connaill, B., Whittaker, S., and Wilbur, S. (1993) Conversations Over Video Conferences: An Evaluation of the Spoken Aspects of Video-Mediated *Communication Human Computer Interaction.* 8: 389-428.
17. Olson, J. S., Olson, G. M. and Meader, D. K. (1995). What Mix of Video and Audio is Useful for Small Groups Doing Remote Real-Time Design Work? *Proceedings of the ACM CHI'95*. 362 - 368
18. Rice, R. E. (1993), Media Appropriateness: Using Social Presence Theory to Compare Traditional and New Organizational Media. *Human Communication Research*, **19**(4), 451-484.
19. Short, J., Williams, E. & Christie, B. (1976*). The Social Psychology of Telecommunications.* Wiley Press, London.
20. Sellen, A., J. (1992). Speech Patterns in Video-Mediated Conversations, in *Proceedings of ACM Conference on Computer-Human Interaction (CHI '92).* 49-59.

21. Vaske, JJ. and Grantham, C.E., (1989). *Socializing the human-computer environment*. Nonwod: Ablex.

22. Whittaker, S. & O-Conaill, B. (1997). The Role of Vision in Face-to-Face and Mediated Communication. In Finn, E. E., Sellen, A.J. and Wilbur, S.B. (Eds.) *Video-Mediated Communication. Mahwah,* NJ: Lawrence Erlbaum. 23-50.

Exploring Communicative Modes for a Virtual Environment: A Social Psychological Analysis

Eva-Lotta Sallnäs
The Royal Institute of Technology,
IPLAB/NADA/KTH
SE-100 44 Stockholm, Sweden
+46 (0) 8 790 66 26
evalotta@nada.kth.se

Anders Hedman
The Royal Institute of Technology,
CID/NADA/KTH
SE-100 44 Stockholm, Sweden
+(0) 8 790 92 83
ahedman@nada.kth.se

ABSTRACT
This paper reports from an experimental study of cooperative work in a collaborative virtual environment (CVE) with sixty test subjects. A method from social psychology, interaction process analysis (IPA), was used for analyzing group dynamics within the CVE in relation to three communicative modes: text-chat, voice, and video conferencing. The study investigated how the communicative mode influenced the quality of the social interaction and how cooperative work in a CVE could be analyzed by using categories of social interaction. It was found that the frequencies of utterances in the voice and video conferencing conditions were significantly larger than in the text-chat condition. The dialogue of the text-chat condition was found to be more task oriented than the dialogues of the voice and video conferencing condition. The social psychological method used was not developed for use within CVE:s, but it was found to work well since the scope of the IPA method is broad and applicable to virtually any small group.

Keywords
Collaborative virtual environment, communicative modes, interaction process analysis, experiment

INTRODUCTION
This paper reports from an experimental study that was conducted in a collaborative desktop virtual environment (CVE). The study aimed at gaining a better understanding of the social dynamics and experiences of people collaborating in such environments as they use different communicative modes. By communicative mode is here meant the means by which people communicate with each other in a CVE. Three different such modes were deployed: text-chat, voice, and video conferencing. In the study presented here the aim was to explore, if the mode of the communication channel influenced the test subjects' task and socioemotional orientation when cooperating in a CVE.

Research on electronic communication has shown how media supporting different modalities vary in capacity for carrying information [20, 7, 18]. Communicating through audio is important when collaborating at a distance and improves both task performance and perceived affordances in comparison to text-chat [15, 20, 21]. In one study [4] various modes of interaction were examined ranging over audio, handwriting, typewriting, video and face-to-face. Results showed that people spoke more than they wrote, and people performed best in terms of time to complete tasks when audio was provided compared to text.

Research has not found such uniform results regarding benefits from using video conferencing. It is usually shown that video does not add significant advantages to task performance compared to audio [22, 1, 17]. Jensen et al [12] examined the nature of cooperation between two players of an electronic text-based game under four different conditions: no communication channel, text-chat, text-to synthesized speech, and audio. They found that immediate forms of communication e.g., audio are important for cooperation and trust in online environments. Jensen points to the need for conducting studies in richer environments with e.g., 3d avatars and researching other aspects of social interaction than cooperation and trust.

BACKGROUND
Social psychology has had considerable impact on research in the field of computer mediated communication–CMC [20, 19, 13, 21]. This paper represents an attempt to explore a method from social psychology as an approach to understanding the social dynamics of CVE:s. It relies to an important extent on Bales Interaction Process Analysis (IPA) taxonomy of social interaction. The IPA taxonomy was not developed with CVE:s in mind, but for physically co-located teams. The findings presented in this paper show that the taxonomy also is useful for analysing social interaction in CVE:s.

Interaction Process Analysis emerged as a way to study group interaction by mapping acts that occur in co-located social interaction [2] and has also been applied in CMC [3, 10, 11]. The IPA evolved through several studies of communication acts in face to face meetings. When interacting, a group is seen as taking part in a problem solving activity consisting of a sequence of acts. These acts give the participant a sense of being part of a joint or shared process. Bales argue that in order to maintain the shared process, people engage in two general categories of communication: task-oriented and socio-emotionally oriented communication. These general categories are also part of a more detailed taxonomy that Bales developed for analyzing verbal behavior.

Experimental studies of small physically co-located groups showed that only a dozen categories were relevant to code in collaborative situations [2]. These categories are meant to be collectively inclusive and mutually exclusive, i.e., every observed communication act can be classified as belonging to exactly one category. The unit to be coded is the smallest segment of verbal or non-verbal behavior to which the observer can assign a classification. Thus complex sentences usually involve multiple code units. Verbal acts within a small-group setting are analyzed into the following low-level categories: give information, give suggestion, give opinion, ask for information, ask for suggestion, ask for opinion, show solidarity, show tension release, show agreement, show disagreement, show tension, show antagonism. From these low-level categories, Bales abstracts four mid-level categories: problem solving attempts, questions, positive reactions, and negative reactions. Finally from these mid-level categories two high-level classes are abstracted: task-oriented communication and socio-emotional communication.

The IPA taxonomy makes it possible to compare groups with each other or a group with itself at different times, i.e., before and after the introduction of some experimental variable. The taxonomy was designed with the aim of capturing aspects of interaction so general that they will be manifest in the behavior of virtually any small group. This latter fact makes it plausible that the taxonomy should also be applicable to group interaction in CVE:s and not only to group interaction in physically co-located groups. Moreover, since the IPA taxonomy is so general with respect to physically co-located groups, it can serve as a real-world yard stick against which to examine small group behavior in CVE:s.

METHOD

Experimental design

A between-groups design was used and the independent variable was a CVE with three different communicative modes:

- text-chat
- voice
- video conferencing

The text-chat was a feature of the CVE-software and it appeared as a window beneath and adjacent to the 3D environment. Since the CVE-software had neither voice communication nor video conferencing as built in features, these latter features had to be arranged externally. In the voice condition, participants communicated through telephones with headsets. In the video conferencing condition, two 21-inch television monitors were used in addition to the telephones with headsets.

Participants

Sixty people participated in the experiment. The participants were students from the Stockholm University and the Royal Institute of Technology along with a few administrative staff members. They were assigned to pairs with a woman and a man in each pair and they each performed a simple decision making task together. Pair members did not know each other prior to the experiment.

The collaborative virtual environment

The CVE was constructed in the ActiveWorlds CVE system (www.activeworlds.com) and had the appearance of a simple exhibition with information points (Figure 1). Access to the CVE was restricted to the participants of the study. Two networked portable PC:s were prepared with clients of the ActiveWorlds CVE system.

These information points had posters with pictures of cars and QuickTime movie clips with model-specific information. Humanlike avatars represented the subjects. Female participants had a traditional female avatar with the following characteristics: brown hair in a ponytail, purple trousers and a T-shirt. The male participants had a traditional male avatar with these attributes: black short hair, white T-shirt, and blue trousers.

Figure 1. The exhibition with information points in the ActiveWorlds environment.

Task

The decision-making task was presented to the pairs of participants as a written scenario:

You have participated in a competition in which you and another participant performed equally well. You therefore have to share the first price: a Volvo car of your choice with

insurance and gas for one year. You will use the car every second month. You will not be able to sell the car. The organizers of the competition now want you to go in to a virtual exhibition and choose the car that you are going to share. You should then decide together which one of you is going to have the car the first month.

The participants were encouraged to discuss this scenario with each other and to interpret it together.

Independent variables

One independent variable was varied: the mode of communication involved in solving the task.

In the video conferencing condition, participants had a TV monitor in front of them and wore headsets enabling them to converse via a telephone connection. The video cameras used to document the interaction served dual purposes and were simultaneously used for the videoconference. The angle of the monitors and the video cameras was about 30 degrees to the participants left side. The video cameras were positioned under the monitor and tilted slightly upwards. This latter arrangement served to minimize the parallax problem, as compared to placing the cameras on top of the monitors. Each participant had the notebook PC with the CVE directly in front of them.

In the voice condition the participants communicated via the telephone connection using headsets and had no video connection.

In the text condition the participants communicated only via a text chat available in the CVE.

Dependent variables

Twenty-seven dialogues out of thirty were transcribed from the video recordings. These dialogues were coded into the twelve categories by the two experimenters independently (Table 1). Detailed descriptions on how to recognize and distinguish the different categories are shown in the listing below. As mentioned earlier, the twelve categories can be classed into four types of acts (A-D). Finally A and B can be classed together as, task-oriented communication, and C and D can be classed together as, socio-emotional communication.

Table 1. Twelve categories are listed with descriptions on how to recognize and distinguish them. These twelve categories can be classed into four acts (A-D). Finally A and B can be classed as task-oriented and C and D can be classed as socio-emotionally oriented communication.

Task-oriented communication:

A. Problem-solving attempts

1. Gives information (GI) = orientation, repeats, clarifies, confirms.

2. Gives suggestion (GS) = direction, implying autonomy for other.

3. Gives opinion (GO) = evaluation, analysis, expresses feeling, wish.

B. Questions

4. Asks for information (AI) = orientation, repetition, confirmation.

5. Asks for suggestion (AS) = direction, possible ways of action.

6. Asks for opinion (AO) = evaluation, analysis, expression of feeling.

Socioemotional communication:

C. Positive reactions

7. Shows solidarity (SS) = raises other's status, gives help, reward.

8. Shows tension release (SR) = jokes, laughs, shows satisfaction.

9. Shows agreement (SA) = shows passive acceptance, understands, concurs, complies.

D. Negative reactions

10. Shows disagreement (SD) = shows passive rejection, formality, withholds help.

11. Shows tension (ST) = asks for help, withdraws out of field

12. Shows antagonism (San) = deflates other's status, defends or asserts self.

Documentation

For all three conditions trials were video recorded using a video mixer in such a way that the participants´ faces, torsos and the interaction in the CVE were captured synchronously on the same tape. The navigation of the participants in the CVE was recorded from the perspective of one subject. Audio was captured using one microphone per participant and the text from the text-chat communication was saved for analysis.

Analysis

The classifications of the coders were compared to ensure that the coding was reliable. The code frequencies of the twelve categories respectively and from the two coders were analyzed with one-way ANOVAs (analysis of variance).

If the data from both coders showed the same results regarding significant differences between the three conditions in the twelve categories then the inter-code reliability was regarded as satisfactory. This analysis showed that both coders had equal results in all twelve categories in all three conditions with one exception. The coders did not have the same results in the category give suggestion (GS), when the text and the video conferencing condition were compared.

RESULTS

A comparison of the frequencies of codes in the twelve categories using one-way ANOVAs in the three conditions show a number of significant differences between the text condition and the voice and video conference condition respectively. However, results reveal no significant differences between the voice and video conferencing conditions with respect to the amount of communication between participants. The mean numbers of code units was almost the same in all categories of both the voice and video condition (Figure 2, Table 2). When performing the task, people were markedly more verbally engaged in the voice and in the video conference conditions in comparison to the text condition. Subjects did not seem as engaged in the discussions in the text condition and those dialogues exhibited numerous misunderstandings.

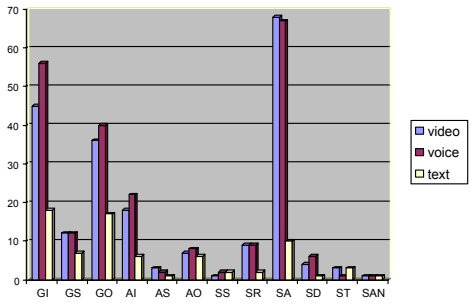

Figure 2. Mean values of frequencies of codes in the twelve categories in the video conference, voice and text condition.

Basic communication categories

One set of categories (GI, GS, GO, AI, AS, AO) is more concerned with task performance. Participants gave information (GI) $F=11$ $p<0.01$, gave opinions (GO) $F=10$ $p<0.01$ and asked for information (AI) $F=24$ $p<0.01$ to a significantly lesser extent in the text condition than in the voice condition. Participants also gave information (GI) $F=8$ $p<0.05$, gave opinions (GO) $F=6$ $p<0.05$ and asked for information (AI) $F=14$ $p<0.01$ to a significantly lesser extent in the text condition than in the video conferencing condition. Participants gave significantly more suggestions (GS) $F=7$ $p<0.05$ in the voice than in the text condition. Note however, that the inter-code reliability was not satisfactory in this single case. Mean values of frequencies of codes in all the twelve categories in the video conference, voice and text condition is shown in table 2.

Table 2. Mean values of frequencies of codes in the twelve categories in the video conference, voice and text condition.

	Video	Voice	Text
GI	M=45	M= 56	M=18
GS	M=12	M= 12	M=7
GO	M=36	M= 40	M=17
AI	M=18	M= 22	M=6
AS	M=3	M= 2	M=1
AO	M=7	M= 8	M=6
SS	M=1	M= 2	M=2
SR	M=9	M= 9	M=2
SA	M=68	M= 67	M=10
SD	M=4	M= 6	M=1
ST	M=3	M= 1	M=3
SAN	M=1	M= 1	M=1

Much of the communication in the voice and video conference conditions concern personal facts such as how large subjects families are, or experiences relevant to the information displayed. Subjects in the voice and video conference conditions also explored the CVE more completely and consequently had more to discuss. Therefore, subjects gave more information (GO) and asked for more information (AI) in the voice and video conference conditions than in the text condition.

Subjects also gave more opinions (GO) in the voice and video conference conditions as they discussed and negotiated more extensively than in the text condition. They often discussed and analyzed the information carefully while at the same time communicating their personal preferences. In the text condition negotiations were more crude.

Another set of categories (SS, SR, SA, SD, ST, San) is concerned with socio-emotional issues. Participants showed tension release (SR) $F=34$ $p<0.01$, showed agreement (SA) $F=34$ $p<0.01$ and showed disagreement (SD) $F=11$ $p<0.01$ to a significantly lesser extent in the text condition as compared to the voice condition. Participants also showed tension release (SR) $F=7$ $p<0.05$, showed agreement (SA) $F=15$ $p<0.01$ and showed disagreement (SD) $F=9$ $p<0.01$ to a significantly lesser extent in the text condition as compared to the video conference condition. Mean values of frequencies of codes in all the twelve categories in the video conference, voice and text condition is shown in table 2.

Subjects showed more agreement (SA) in the video conference and voice conditions that in the text condition. This indicates that it was easier to give feedback in the voice and video conference condition than in the text condition. The difference between conditions with respect to the category show disagreement (SD), reflects that subjects negotiated more in the video conference and voice conditions. This category reveals that it was easier to give feedback about misunderstandings and thus to maintain a common ground in the voice and video conference [6].

Subjects showed tension releasing (SR) behavior more in the voice and video conference conditions than in the text condition. They were joking more and often communicated satisfaction with individual or cooperative behavior. There were examples of tension releasing behavior in the text condition also. Smilies (emotionally expressive symbols created with standard keyboard characters) were sometimes used to communicate jokes or feelings of satisfaction.

Communication acts

In the next step the dialogue was classed into four categories of acts: problem solving, questions, positive reactions and negative reactions in the three conditions and analyzed in a relative manner (Table 3). This relative way of comparing the voice, video conference and text conditions revealed that the main focus in the dialogue was, in all conditions, on problem solving, but most evidently so in the text condition. Moreover, in the text condition, more questions were asked than in the voice and the video conferencing conditions. In the video and voice condition more emphasis was put on conveying positive reactions than in the text condition. The results also show that little verbal communication was about showing negative reactions in any of the conditions.

Table 3. Percentage of acts of problem solving, questions, positive reactions and negative reactions in the video, voice and text-chat conditions.

	Video	Voice	Text
Problem solving	46%	48%	59%
Questions	14%	14%	18%
Positive reaction	38%	35%	20%
Negative reactions	2%	3%	3%

Taskoriented vs. socioemotional communication

The findings were analyzed one step further and the categories were classed together as, task-oriented communication or socioemotional communication (Table 4).

Table 4. Percent of task-oriented and socio-emotionally oriented communication in the video conferencing , voice and text conditions.

	Video	Voice	Text
Task-oriented	60%	62%	77%
Socio-emotional	40%	38%	23%

Communication in all conditions was more concerned with the task than with socio-emotional aspects, most markedly in the text-chat condition. The analysis also reveals that the communication was more socio-emotionally oriented in the video conferencing and voice conditions than in the text condition.

In this phase the analysis of the results shows a tendency that the dialogue in the video condition is more socio-emotionally oriented than the dialogue in the voice condition. The voice condition on the other hand is slightly more task oriented than the video condition.

DISCUSSION

This study has shown that IPA can be successfully applied to the analysis of small group dynamics in a CVE. The categorization was experienced as intuitive and sound by the experimenters and yielded interesting results.

The results presented here are in support of the claim made by Jensen [12] that the richness of the communication channel is an important determinant for the quality of online cooperative work. Communication in the text condition was terse and did not allow the subjects to engage themselves in the more varied verbal behavior evident in the other two conditions. The cost of communicating by using text-chat appeared too great to allow for elaborate discussions or substantial socializing behavior.

The value of voice communication was almost on par with video conferencing. This finding is also supported by earlier work [5, 8, 16]. In comparison to voice, video conferencing was not found to enhance the quality of cooperative work or increase the amount of socializing to a large extent. However, the video channel was used by the subjects in certain situations. The video channel was used when: there were long pauses in the dialogue and both subjects investigated information, when navigation problems occurred, during greetings and during discussions prior to important decision making. Results from the high-level analysis of the categorization into two classes, task- and socio-emotionally oriented communication, do show a tendency that it do matters which of these modes that are used for communication in a CVE.

The reason why more than half of the dialogue in the text condition was focused on problem solving might be that, because communicating in this mode was harder, most of the effort was put into actually solving the task.

The deployment of Bales taxonomy for classifying small group behavior proved to be a viable approach in analyzing the interaction within the CVE under the three conditions. However, as Bales [2] points out the IPA taxonomy was designed with the aim of capturing aspects of interaction so general that they will be manifest in the behavior of virtually any small group.

Issues regarding the specific content and context in a certain case of social interaction, as for example in this specific exhibition in a CVE is not addressed by IPA. From the perspective of Bales [2] a content analysis would be important to conduct in order to obtain complementary information about the specifics of a certain case. The next step would therefore be do conduct a content analysis of the transcripts and video recordings from the study presented here.

ACKNOWLEDGMENTS

Advice and valuable comments from Kerstin Severinson-Eklundh are gratefully acknowledged.

REFERENCES

1. Anderson, A., Newlands, A., Mullen, J., Flemming, A. M., Doherty-Sneddon, G. and Van der Velden, J. (1996). Impact of Video-Mediated Communication on Simulated Service Encounters. *Interacting with Computers*. 8: 193-206.
2. Bales, R. F. (1976). *Interaction process analysis: a method for the study of small groups*. Chicago : Univ. of Chicago Press.
3. Carletta, J., Anderson, A. H., & McEwan, R. (2000). The effects of multimedia communication technology on non-collocated teams: a case study. *Ergonomics*, **43**(8), 1237-1251.
4. Chapanis, A. (1975). Interactive Human Communication. *Scientific American*. 232: 36-42.
5. Chapanis, A., Ochsman, R. B., Parrish, R. N. & Weeks, G.D. (1972) Studies in interactive communication: The behavior of teams during co-operative problem solving. *Human Factors*, 14: 487-509.
6. Clark, H. H., Brennan, S. E., (1991) Grounding in communication. In Perspectives on Socially Shared Cognition. Resnick, L. B., Levine, J. M., and Teasley, S. D. Eds. *American Psychological Association, Washington*. 127-149.
7. Daft, R. L. & Lengel, R. (1986). Organizational Information requirements, media richness, and structural design. *Management Science*. 32: 554-571.
8. Egido, C. (1990) Teleconferencing as a technology to support cooperative work. In Intellectual Teamwork. J. Galegher, R. Kraut, C. Egido, Eds. *Lawrence Erlbaum, Hillsdale, NJ*. 351-371.
9. Fraser, M., Benford, F., Hindmarsh, J., & Heath, C. (1999). Supporting Awareness and Interaction through Collaborative Virtual Interfaces In *Proceedings of UIST '99*, Asheville, NC, 1999. 27-36.
10. Hiltz, S.R., Johnson, K., & Agle, G. (1978). Replicating Bales' problem solving experiments on a computerized conference: *A pilot study (Research Report No. 8)*. Newark: New Jersey Institute of Technology, Computerized Conferencing and Communications Center.
11. Hiltz, S.R., Johnson, K. & Rabke, A.M. (1980). The process of communication in face to face vs. computerized conferences: A controlled experiment using Bales interaction process analysis. In N. K. Sondheimer (Ed.), *The 18th Annual Meeting of the Association for Computational Linguistics and Parasession on Topics in Interactive Discourse: Proceedings of the Conference.* Philadelphia: University of Pennsylvania.
12. Jensen, C., Farnham, S. D., Drucker, S. M. & Kollock, P. (2000). The Effect of Communication Modality on Cooperation in Online Environments. In *Proceedings of the ACM Conference on Computer-Human Interaction (CHI '00),* Amsterdam. 470-477.
13. Kraut, R. E., Fish, R. S., Root, R.W., & Chalfonte, B.L. (1990). Informal communication in organizations: Form, function, and technology. In S. Oskamp & S. Spacapan (Eds)., *Human Reactions to Technology: The Claremont Symposium on Applied Social Psychology. Beverly Hills, CA: Sage Publications*. 145-199.
14. Matsuura, N., Fujino, G., Okada, K. and Matsushita, Y. (1993). An Approach to Encounters and Interaction in a Virtual Environment. *Proceedings of ACM Computer Science Conference*. Association for Computing Machinery. 298-303.
15. Ochsman, R.B. and Chapanis, A. (1974). The Effects of 10 Communication Modes on the Behavior of Teams During Co-operative Problem-solving, International *Journal of Man-Machine Studies*. 6: 579-619.
16. O'Connaill, B., Whittaker, S., and Wilbur, S. (1993) Conversations Over Video Conferences: An Evaluation of the Spoken Aspects of Video-Mediated *Communication Human Computer Interaction.* 8: 389-428.
17. Olson, J. S., Olson, G. M. and Meader, D. K. (1995). What Mix of Video and Audio is Useful for Small Groups Doing Remote Real-Time Design Work? *Proceedings of the ACM CHI'95*. 362 - 368
18. Rice, R. E. (1993), Media Appropriateness: Using Social Presence Theory to Compare Traditional and New Organizational Media. *Human Communication Research*, **19**(4), 451-484.
19. Short, J., Williams, E. & Christie, B. (1976*). The Social Psychology of Telecommunications.* Wiley Press, London.
20. Sellen, A., J. (1992). Speech Patterns in Video-Mediated Conversations, in *Proceedings of ACM Conference on Computer-Human Interaction (CHI '92).* 49-59.

21. Vaske, JJ. and Grantham, C.E., (1989). *Socializing the human-computer environment*. Nonwod: Ablex.

22. Whittaker, S. & O-Conaill, B. (1997). The Role of Vision in Face-to-Face and Mediated Communication. In Finn, E. E., Sellen, A.J. and Wilbur, S.B. (Eds.) *Video-Mediated Communication. Mahwah,* NJ: Lawrence Erlbaum. 23-50.