BEHAVIOUR & INFORMATION TECHNOLOGY, 2000, VOL. 19, NO. 5, 315–327

Knowledge acquisition in ecological product design: the effects of computer-mediated communication and elicitation method

JÜRGEN SAUER, SIMONE SCHRAMME and BRUNO RÜTTINGER

Institute of Psychology, Darmstadt University of Technology, Hochschulstr. 1, D-64289 Darmstadt, Germany; e-mail: sauer@psychologie.tu-darmstadt.de

Abstract. This article presents a study that examines multiple effects of using different means of computer-mediated communication and knowledge elicitation methods during a product design process. The experimental task involved a typical scenario in product design, in which a knowledge engineer consults two experts to generate knowledge about a design issue. Employing a 3×2 between-subjects design, three conference types (face-to-face, computer, multimedia) and two knowledge elicitation methods (structured interview, network technique) were compared. One hundred and eight participants took part in the study. They were assigned to 36 groups of three. Quantitative and qualitative performance data were collected and the group processes with the IPA method analysed. The results showed that the computer conference group was generally more productive than the two other groups during the conference. However, participants were unable to maintain their higher performance levels in a later task where the conference results had to be edited before being fed into an expert system. As expected, the computer conference group showed the lowest socio-emotional content during interaction. The network technique was largely more productive than structured interviewing, though it was more time-consuming. Furthermore, the findings suggested that both the network technique and computer conferencing achieved their higher productivity in knowledge generation only at the cost of information processing depths, resulting in poorer performance for subsequent transfer activities. The results are discussed with regard to organizational choice in managing conferences of this kind.

1. Introduction

In the context of product design, teams have been widely used to take advantage of the wider knowledge base of groups compared to individuals, and the complexity of the process renders this task virtually unmanageable by a single person (Frankenberger 1997). In order to manage the complex requirements of the design process, a core design team is formed to direct the various stages of the process. In addition to the core team, experts are consulted concerning various aspects of the design activities, as this allows the widening of the knowledge base of the design team (Scott *et al.* 1991). The experts usually come from very different professional backgrounds, such as marketing, materials science, production processes, control or environmental science. Although the experts are not part of the core design team and are involved on a non-permanent basis, they play a very important role in supporting the core group (Ehrlenspiel 1995).

As designers have to consult experts who may work at a distant location, the use of computer-mediated communication (CMC) is not uncommon to save costs and time of travelling. While there has been an increasing proliferation of systems supporting CMC in a wide range of areas (e.g. mailing, project management, conferencing, diary management; for an overview see, Johansen 1988), the effects of using CMC are very context specific and require assessment of its multiple effects on important variables, such as performance parameters, participant satisfaction, socio-emotional content (McCarthy and Monk 1994).

1.1 *Computer-mediated communication*

This article is concerned with holding team meetings in a conference setting, which is an increasingly used application for CMC instruments. While face-to-face communication is still of high importance for confer-

Behaviour & Information Technology
ISSN 0144-929X print/ISSN 1362-3001 online © 2000 Taylor & Francis Ltd
http://www.tandf.co.uk/journals

ences, with changes in work organization (e.g. telework, outsourcing, globally operating organizations) others forms of communication have shown to be viable alternatives. There are several alternatives available to face-to-face communication, such as video conferencing and computer conferencing (Johansen 1988). They differ with regard to the fidelity with which they simulate aspects of face-to-face communication. For the purpose of drawing a comparison between means of communication, they may be rated on three dimensions: media richness (or communication bandwidth), geographical scope and time delay in communication (Barua *et al.* 1997). This helps produce a utility profile for each means of communication, which also takes into account the overall circumstances and the task for which it is being used. For example, face-to-face meetings enjoy a high level of media richness but suffer from low geographical scope (i.e. groups must not be spatially distributed).

Communication media also differ with regard to the underlying social context cues during communication (Straus and McGrath 1994). While multimedia conferences (e.g. with a PC-based video link) model much of the bandwidth of communication channels used in face-to-face communication (auditory and visual information of different kind, non- and para-verbal cues), in the case of computer conferences this is reduced to text-based information. However, reduced bandwidth does not necessarily have to be a disadvantage. As some people have argued, computer conferences have a number of benefits. For example, the participation rate is more equally balanced (i.e. the discussion is not controlled by dominant group members) than in face-to-face meetings (Sproull and Kiesler 1986, Walther 1995, Straus 1996). It has been found that a more balanced participation rate generally makes better use of group resources (Rüttinger *et al.* 1994). Furthermore, it has been argued that in computer conferences, differences in status do not affect communication patterns to the same degree as during face-to-face communication (Sproull and Kiesler 1986). However, the findings are not unequivocal since other research suggested that status did affect the participation and influence of group members even during computer conferencing (Weisband *et al.* 1995).

Multimedia as a more recent CMC tool uses several media (e.g. video, audio, text, animation, graphics) that are integrated in a single communication environment. Due to its multiple mode features that produce a higher level of media (or interaction) richness, it resembles in some ways face-to-face communication, though the quality and quantity of non-verbal cues is lower. Furthermore, additional demands accrue because of the multiple task nature of the multimedia environment, which requires integration of different task elements. Due to its more recent emergence, there has been little research into multimedia applications, though one may draw upon work examining video conferences as the closest analogue (e.g. Finn *et al.* 1997).

CMC research has also pointed at differences in task-orientation and relationship-orientation in communication patterns. In particular, computer conferencing has been associated with higher task-orientation and lower socio-emotional content (Rüttinger *et al.* 1994). Work by Rice and Love (1987) also confirmed the dominance of task-orientation over socio-emotional content in CMC, though they argue that socio-emotional messages are much more prevalent than commonly assumed. Furthermore, there is evidence for a time-based effect. Socio-emotional orientation increases in CMC if the group process evolves over a longer period of time, leading to a diminishing difference between face-to-face and computer conferencing (Walther and Burgoor 1992, Walther 1995).

1.2. *Knowledge elicitation*

While the effects of CMC have attracted considerable research interest, few studies have examined the use of elicitation techniques in the context of product design. There is a wide range of techniques available for knowledge elicitation, such as structured interviews, simulations, sorting tasks and role play (see Cordingley 1989). One may distinguish between different types of knowledge elicitation techniques, based on the source of the information or the degree of formalization of the elicitation technique. A classification system by Konradt (1992) distinguishes between three types: inquiry techniques, observation, and testing. Inquiry refers to direct person-centred data acquisition methods, such as interviews, focus groups and questionnaires. Observation is concerned with techniques such as videotaping and verbal protocols. The category testing refers to techniques that use structured material that is to be administered with standardized instructions. Examples of these are repertory grids, network and charting techniques. Each of the techniques naturally has its strengths and weaknesses, which means its effectiveness is largely dependent on the context in which it is employed (see Bainbridge 1979, Cordingley 1989).

Network techniques (see Kirwan and Ainsworth 1992) support the organization and structuring of knowledge, with less variation in performance as a result of the abilities of the knowledge engineer. The elicited knowledge is well structured, which considerably facilitates subsequent data analysis (see method for a

more detailed discussion). Interviewing, in contrast, requires considerable time and effort during the process of data analysis. It is a flexible method that is applicable to a wide range of subject areas. Furthermore, it requires less specialized training for the knowledge engineer than network techniques.

1.3. *The present study*

So far CMC research has primarily focused on 'process' criteria (i.e. participation rate, interaction patterns) while the use of performance indicators has usually been limited to measuring conference duration. Furthermore, the task scenarios used had generally little in common with typical applications of CMC in the real world.

Against this background, the article aims to present empirical data about the effects of using different forms of CMC in a realistic team-based task environment, modelled on the product design process in extended design teams. In this study, a range of performance measures and process variables were employed to gain a fuller picture of the multiple effects of using different conference types. The multi-level analysis examined various aspects of conference effectiveness as well as group interaction processes (see McCarthy and Monk 1994). A person-centred and a method-centred knowledge elicitation technique was chosen for the experimental work to address the question of how the match between conference type and knowledge elicitation can be improved.

It was expected that computer conferencing would be more productive than the other two conference types because it is more task-focused due to its small communication bandwidth (see also Rüttinger *et al.* 1994). This should be advantageous for the task used in this study, which does not require social context cues. These would be needed for tasks with higher ambivalence levels (Straus and McGrath 1994). The low media richness of computer conferences should also be reflected in a lower socio-emotional content of communication patterns (i.e. positive as well as negative ones), which should be demonstrated by the IPA analysis (Bales 1970). Due to its complex technical environment, the multimedia conference was expected to show poorer performance than the face-to-face conference. The socio-emotional content was expected to be highest in face-to-face conferences since it allowed the most immediate form of communication. Finally, it was predicted that a network technique would be the more effective elicitation method since it provides a better structure to the problem-solving activity (see Hacker and Jilge 1993).

2. Method

2.1. *Design*

A 3×2 between-subjects design was employed, with 36 teams of three members each. With the unit of analysis being the team, this meant that there were six cases in each cell. The independent variable conference type was varied at three levels: face-to-face conference (FTF), computer conference (CC) and multimedia conference (MM). The second independent variable knowledge elicitation technique was varied at two levels: structured interview (SI) versus HSLT (a network technique which is described below).

2.1.1. *Conference type*: During FTF participants sat around a table, allowing for direct communication. During CC, participants were seated in separate rooms, using Silicon Graphics Workstations (Indys) for communication. Using different rooms was considered a better simulation of this kind of conference than merely separating participants by screens while being situated in the same room, as it has been frequently done in laboratory-based experiments. Being placed in separate rooms, too, and using the same work stations, participants in the MM condition communicated by means of a conference system, called 'Inperson' (Silicon Graphics). There were video pictures of the other participants presented on screen while auditory information was received by headphones. Furthermore, the system provided an embedded whiteboard for reading and writing messages.

2.1.2. *Knowledge elicitation technique*: The effectiveness of two knowledge elicitation techniques were compared as a function of the conference type employed. While in practice interviewing has been very widely used for the purpose of knowledge engineering for design (Scott *et al.* 1991), attempts have been made to use a more structured approach to knowledge elicitation (see Hacker and Jilge 1993). A further difficulty is that the success of SI is strongly dependent on the preparation made by knowledge engineers prior to data acquisition and on their ability to ask the 'right' questions. Thus designers have asked for an elicitation technique that produces more structured material. For this purpose, a network technique, called HSLT (Heidelberger Struktur-Lege-Technik) was chosen, that has been widely used in the German psychological community (Scheele and Gröben 1979). As HSLT is less known in the English research literature, it is briefly described below.

This method aims to create an explicit structure of the knowledge that an expert has about a particular

area. In the first phase, the principal concepts are written on cards. In the second phase, the concepts are related to each other by using a number of rules that define the relationship between two concepts (e.g. 'A is a part of B', 'C is an example of D', 'E is a feature of F'). The rules that are used during a session may be selected on the basis of the kind of knowledge to be elicited. HSLT aims to produce a hierarchical semantic network as used in cognitive psychology for the representation of human knowledge (e.g. Eysenck and Keane 1995). These networks distinguish between 'objects' and 'relations'. Objects refer to animate and inanimate things (e.g. bear, hut, tree) while relations refer to the link between objects and their characteristics (e.g. subordinate, an example of, a feature of). As one is more concerned with abstract terms, one will refer to objects as 'concepts'. An example of a network structure created by HSLT is presented in figure 1. The task in this example was to design an environmentally-friendly PC. The structure shows a number of components of a PC, which were supplemented by features that participants considered to be relevant for environmental friendliness of the PC.

2.2. *Participants*

A total of 108 participants took part in this study. They were all male students of Darmstadt University of Technology. Their ages ranged from 20–36 years (mean: 24.0 years). All of them were reading technical subjects at the university (electronic and mechanical engineering, computer science, etc.). To fulfil selection requirements, participants needed to have some knowledge in the specific subject areas. This was verified by using a basic selection procedure comprising an interview and a short written test measuring knowledge of ecological issues in computer design.

The test comprised six multiple-choice items (e.g. 'Which one of the following components of a PC consumes most energy?'). The results of the test were the basis for assigning the participants the experimental role of a novice or an expert. While in general a minimum score of three was required for participants to be considered an expert, a subjective assessment of the candidate based on the interview session was also used to decide what role they should play in the experiment. However, if a prospective participant did very poorly on

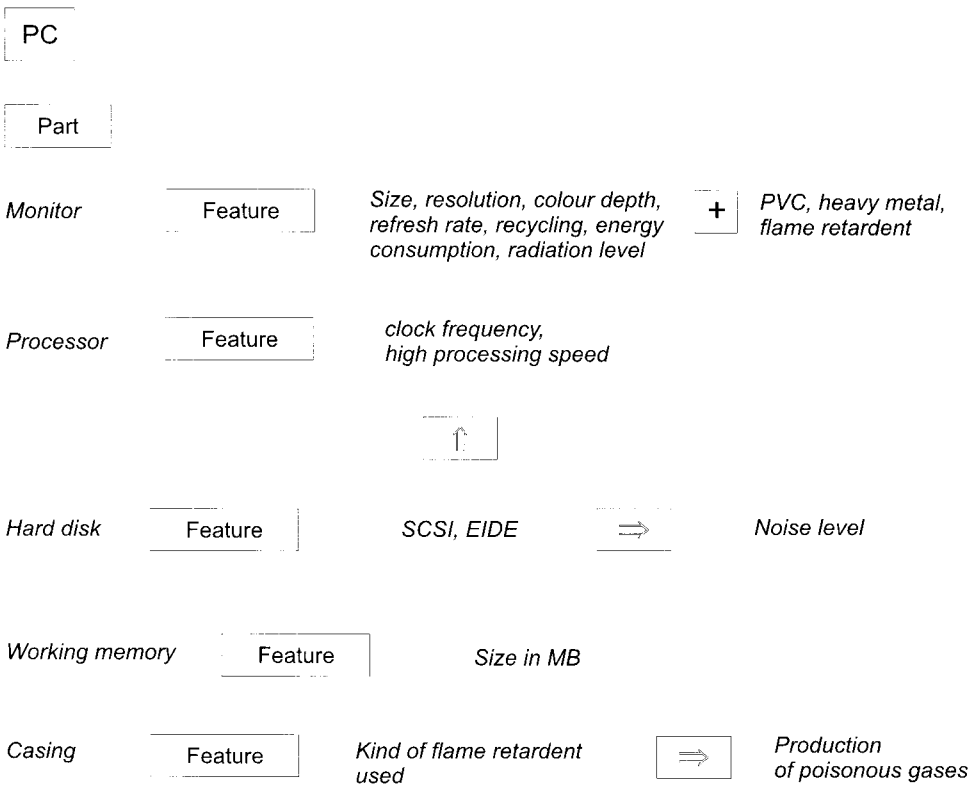


Figure 1. Example of a network structure created by HSLT: Relations are presented in boxes (e.g. is a feature of); concepts are presented in italics (e.g. hard disk); arrows symbolize 'to affect'; plus-sign symbolises 'and'.

the test (only 0 – 1 items correctly answered), he was not allowed to take part at all, as his knowledge base was considered insufficient.

2.3. *Dependent variables*

The experimental sessions in the FTF and MM conditions were videotaped. For the CC condition, the on-screen information was logged into a results file at two-minute intervals, representing the experimental protocol.

Four central performance variables were measured, distinguishing between two quantitative and two qualitative performance measures.

'Conference duration' was an indicator of the time requirements of the knowledge elicitation process. The productivity of the knowledge elicitation session was measured by a 'knowledge generation index', which ıdicated the number of generated concepts per time unit. Both variables were quantitative performance measures. Their strengths is the objectivity with which aspects of performance are evaluated but they are less useful for the assessment of overall performance.

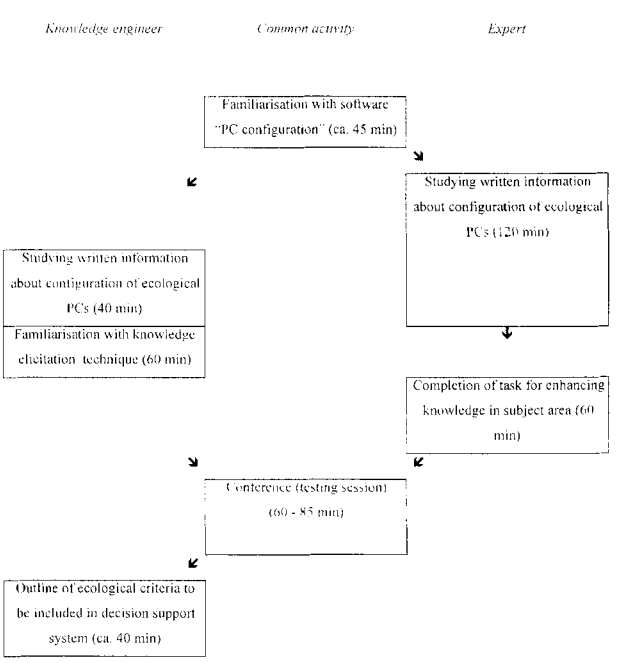
Therefore, two qualitative performance indicators were taken as a supplement to gain a more complete picture. First, a qualitative 'evaluation of the conference outcome' was carried out by two expert raters, using a scoring system based on a 6-point scale. Second, the 'data for expert system input' were also evaluated by the raters employing the same scale. A more detailed description of each performance measure is given in the results section.

Of no lesser importance was the collection of data about the interaction process during the conference since this may be indicative of communication efficiency and group climate. This was achieved by using the interaction process analysis (IPA) method, developed by Bales (1970), which is one of the most widely-used ıethods for analysing group processes. Based on a set of 12 categories, the IPA method aims to collect a number of indices about interaction patterns in groups. It emphasises the process of communication rather than the content of the message. For the purpose of this study, it allowed us to test the hypotheses (see above) with regard to the socio-emotional content of the three conference types. The frequency distribution of messages across categories was examined, a central aspect of the many analyses offered by the IPA method.

2.4. *Training*

Considerable training was necessary before taking part in the experimental session. Training took place over two sessions and lasted for up to 3.75 hours. The different tasks of the group members required that novices (i.e. knowledge engineers) and experts received in part different training regimes, which are described below. Of the total number of participants, 36 were trained as knowledge engineers, and 72 as experts. An overview of the training regimes may be found in table 1. The first part was identical for both groups, in which they were familiarized with a purpose-built software, called 'PC Configuration'. Built for the purpose of the overall research project, this is a computer-supported information system, designed to provide decision support to individuals for purchasing computers. 'PC Configuration' provides prices of hardware components and their compatibility, which allows the decision-maker to select a configuration that fulfils functional requirements (e.g. job demands) and is cost-effective at the same time. However, 'PC Configuration' does not contain any information about the ecological properties of hardware components. Once the participants were familiar with 'PC Configuration', they moved on to acquiring basic information about the subject area of ecological aspects of 'PC configuration'. The information acquisition process differed however in length and content between knowledge engineers and experts, as shown in table 1. All participants received the same training, independently of the communication medium used (i.e. FTF were also trained on 'PC configuration').

Table 1. Structure of training and testing session.



2.4.1. *Knowledge engineer*: With training taking place in groups of two, knowledge engineers acquired knowledge of the subject area (i.e. configuration of ecological PCs) by reading two articles from computer journals. This was to reduce possible knowledge differences between participants. Furthermore, it provided participants with a sufficient knowledge base that would enable them to apply a knowledge elicitation technique successfully. Following the basic training session, the knowledge engineer was made familiar with his respective knowledge elicitation method.

For the structured interview technique, the task was to develop the interview schedule. To gain some familiarity with the technique, knowledge engineers were asked to practise it in the context of football. After some practice has been gained, knowledge engineers in groups of two developed the interview schedule for the testing session. For HSLT, the training sessions were of a similar structure. After having been given some information about the rationale of the technique, knowledge engineers also practised HSLT in the context of football.

2.4.2. *Experts*: With experts also being trained in groups of two, they were asked to carry out a document analysis of relevant material about ecological issues in computer design. They were given 12 articles from computer journals to enhance their knowledge of the subject area. Altogether, 120 minutes were allocated for this activity, that is, three times longer than the knowledge engineers were given time to read.

In the following session, the experts were given the opportunity to practice the knowledge elicitation process with a knowledge engineer. During task completion in the training sessions, all participants already worked with the technical task environment to be used in their experimental condition (i.e. CC or MM system) to obtain a satisfactory degree of familiarity for the experimental sessions.

2.5. *Scenario*

The experimental work was based on a typical scenario in product design, in which the designer consulted two experts about design issues in a conference setting. The task was to produce a structure of knowledge needed for environmentally-friendly computer configuration. The knowledge was then to be arranged such that it could be fed into the database of a knowledge-based expert system. No time limit was given for this task, though participants were told to work on the task for about an hour at least. This was an important point because Walther (1992) argued that some artefacts in CMC research emerged because of experiments being cut off too early, which did not allow for more relational communication patterns to occur in the generally slower CMC.

When the conference finished, the experts had completed their part of the experiment (see table 1). The knowledge engineer was then asked to make suggestions for the enhancement of a computer-supported information system by including ecological criteria in 'PC configuration'. The knowledge engineer was told that these recommendations were subsequently to be implemented in the next version of 'PC Configuration', though this step was not part of their task.

3. Results

3.1. *Group performance*

3.1.1. *Conference duration*: The duration of the conference was an indicator of the time requirements of the knowledge elicitation process (see figure 2). The results showed that multimedia conferences took longest to complete (81.7 minutes), followed by computer conferences (77.8 minutes) and face-to-face sessions (62.3 minutes). While this difference was statistically significant ($F = 4.36$; $df = 2,30$; $p < 0.05$), the more interesting effect was the interaction between media and elicitation method ($F = 5.24$; $df = 2,30$; $p < 0.05$). As figure 1 shows, there was little difference between elicitation method for FTF and CC but a strong increase in duration for MM conferences using the HSLT method. A main effect of the elicitation method was also observed ($F = 8.18$; $df = 1,30$; $p < 0.001$), with HSLT

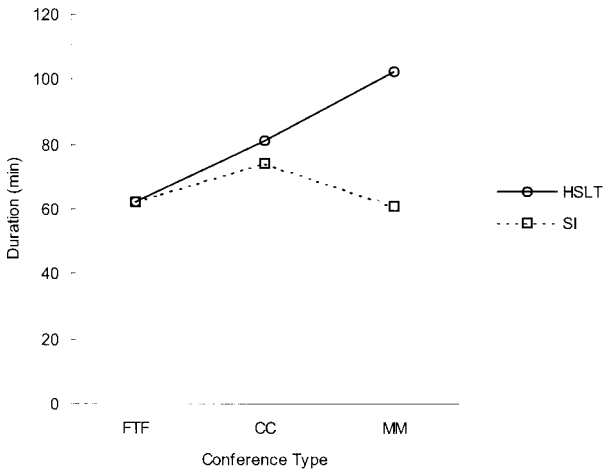


Figure 2. Conference duration (minutes) as a function of medium and elicitation method.

requiring more time than the interviews (82.0 versus 65.8 minutes).

3.1.2. *Knowledge generation*: While conference duration is an indicator of the time commitments required in different conditions, it does not provide a productivity measure of the session. Thus the number of concepts were examined (e.g. low radiation, recyclability, ecological label) that were elicited during the simulated conference session. To control for conference duration as a covariate, an efficacy index was calculated that comprised the number of concepts generated per time unit (minute). However, the analysis did not show any significant difference between elicitation methods ($F<1$), though the HSLT produced somewhat more concepts (0.30/minute) than the structured interviews (0.23/minute). No other effects were found (all $F<1$). The number of relations was then examined, defining the relationship between concepts and their properties. Again, an efficacy index was calculated. Here, a very strong effect of elicitation method was found ($F = 100.8$; $df = 1,30$; $p<0.001$), with HSLT being more than six times as productive as the interviews for that measure (see figure 3). A main effect of conference medium was also observed ($F = 8.98$; $df = 2,30$; $p<0.001$). Computer conferences were most productive, followed by FTF conferences and MM. No significant interaction was observed.

3.1.3. *Qualitative evaluation of conference outcome*: As the analysis of quantitative performance indicators cannot entirely capture the complexity of task performance, a qualitative evaluation of the conference outcome was also carried out. This was done by two raters with an engineering background, with both having considerable knowledge of ecological issues in computer design. On the basis of written documents or computer logs, the two judges assessed the quality of the generated solutions (the judges were not informed about the experimental condition each participant was assigned to). The inter-rater reliability coefficient calculated was satisfactory (Cohen's Kappa = 0.76). For performance assessment, the raters used a scoring system based on German school grades, ranging from 1 (very good) to 6 (very poor).

The results of this analysis is presented in figure 4. It shows a main effect of elicitation method, with HSLT producing better results than the interviewing technique for all conference types ($F = 8.25$; $df = 1,30$; $p<0.005$). The quality of the conference also varied as a function of medium ($F = 3.54$; $df = 2,30$; $p<0.05$). Computer conferences were considered to produce the best results while the scores of MM and FTF conferences indicated poorer outcomes (LSD-test: $p<0.05$). Furthermore, an interaction was recorded, with MM conferences showing a big difference in quality between HSLT and interviews ($F = 3.54$; $df = 2,30$; $p<0.05$). The poorest result by far occurred in the MM condition using interviews while the same medium produced the best score under HSLT.

3.1.4. *Qualitative evaluation of processed data for expert system*: The final task of the knowledge engineer was to aggregate the information gained from the expert discussion and to make suggestions for the enhancement of the decision support system 'PC Configuration' by

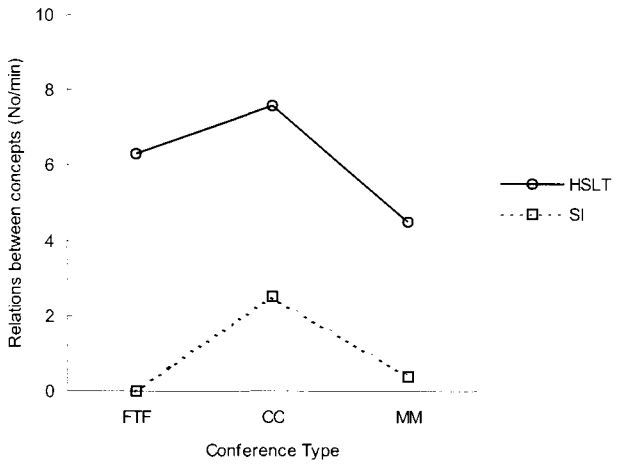


Figure 3. Efficacy index for generation of relations (no of generated relations between concepts per minute) as a function of medium and elicitation method.

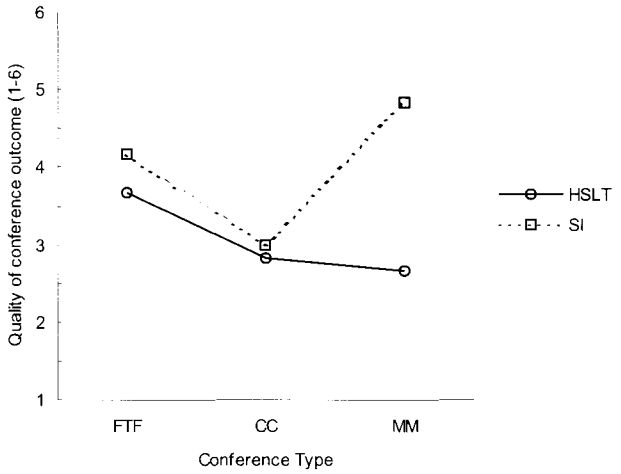


Figure 4. Score of qualitative evaluation of conference results (1 – 6) as a function of medium and elicitation method (1 = very good; 6 = very poor).

incorporating ecological issues. To examine performance on this task, our two judges assessed the quality of the recommendations for the decision support system, using the same scoring system described above. Again, the inter-rater reliability coefficient was judged satisfactory (Cohen's Kappa = 0.79).

As the results in figure 5 show, the pattern is considerably different from the qualitative evaluation of the conference. Again, FTF and MM conferences had a fairly similar score while the CC condition deviated from it but, this time, in a different direction. Knowledge engineers using the CC medium had difficulties in transforming the demonstrated good conference outcome into data that were to be fed into the expert system, in particular, when using structured interviews. However, analysis of variance just failed to confirm the statistical significance of this observation ($F = 2.81$; $df = 2,30$; $p > 0.05$). The results also showed a better performance with the interviewing technique than with HSLT for two of the conditions (FTF and MM). Although this observation cannot be confirmed by a statistically significant interaction ($F < 1$), the pattern is in contrast to other performance indicators, of which most showed better performance for HSLT. No main effect of elicitation method occurred ($F < 1$).

3.1.5. *Performance stability index*: The analysis suggested a change of performance at the second assessment (expert system input) compared to the first assessment (conference outcome). This was examined by calculating a difference index ($score_{conference\ outcome} - score_{expert\ system\ input}$). The analysis showed that the CC group showed a clear loss of performance (-0.9) while the MM group and the FTF group showed improvements ($+1.0$ and $+1.1$). This difference was statistically significant ($F = 8.13$; $df = 2,30$; $p < 0.005$). A significant change was also found for knowledge elicitation method ($F = 4.76$; $df = 1,30$; $p < 0.05$). Here the interviewing method showed relative improvements over the two points of measurement ($+0.8$) while little change was observed for the HSLT (-0.1). No significant interaction was recorded.

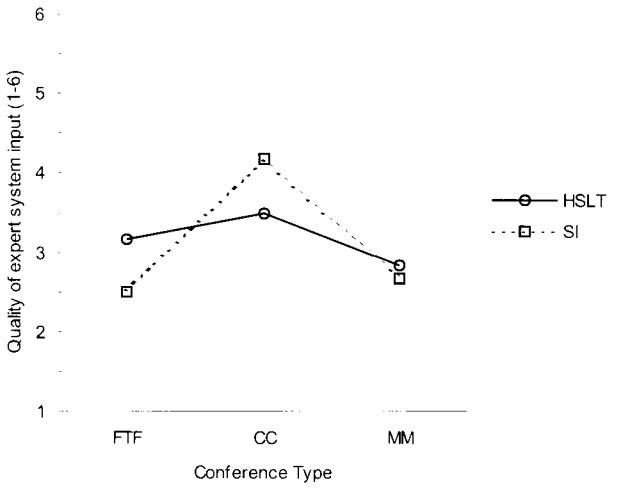


Figure 5. Score of qualitative evaluation of data for expert system (1–6) as a function of medium and elicitation method (1 = very good; 6 = very poor).

3.2. *Interaction process analysis*

As a complement to the range of performance measures, the interaction processes that took place during the conference sessions were also examined by using the IPA method (Bales 1970). To reduce complexity of data presentation, the 12 categories were summarized into the four superordinate categories provided by Bales: Positive (and mixed) actions, negative (and mixed) actions, attempted answers and questions. Frequency data from an IPA analysis are often transformed into percentages to compare them with Bales' estimated norms. Before making such a comparison, a two-way analysis of variance on the frequency data was carried out. The total frequencies were combined with conference duration to an index of frequency per time unit (i.e. number of interactions per minute).

Overall, conference discussions were dominated by task-related interactions (i.e. answers and questions) compared to socio-emotional interactions (i.e. positive and negative actions). Furthermore, positive actions were clearly more frequent than negative actions. As figure 6(a–d) shows, most interactions fell into the categories 'attempted answers' (7.7/minute) and 'positive actions' (3.3/minute). Considerably fewer recordings were made for 'negative actions' (0.7/minute) and 'questions' (0.6/minute). Furthermore, the data suggested that FTF and MM conferences were remarkably similar with regard to the interaction process while CC differed considerably from the two.

Due to the reduced communication bandwidth in CC, it was not surprising that overall frequencies of interactions were lower for this condition than for the two others. This was statistically significant for all four categories (all $p < 0.001$). No differences between MM and FTF were found for any of the categories. Across all categories, more interactions between conference participants were observed in HSLT conditions than during structured interviews, though these were only significant for positive actions ($F = 4.88$; $df = 1,29$; $p < 0.05$), answers ($F = 9.75$; $df = 1,29$; $p < 0.005$) and negative

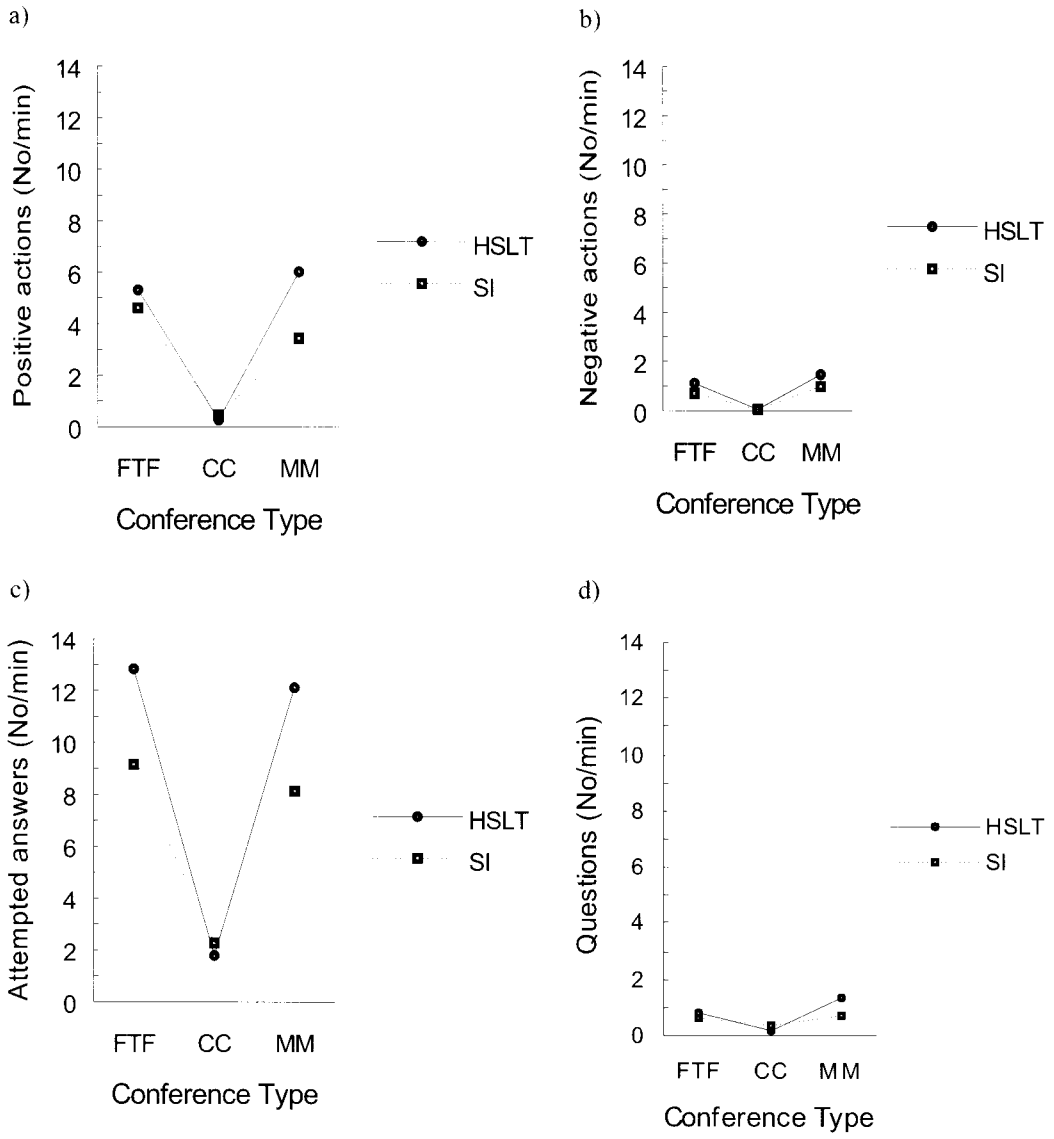


Figure 6. (a–d) Data of interaction process analysis as a function of medium and elicitation method.

actions ($F = 4.30$; $df = 1,29$; $p < 0.05$). Although no significant interactions were observed between conference type and elicitation technique, visual inspection of the data indicated that the difference between HSLT and interviewing only emerged for MM and FTF but virtually disappeared during CC.

In order to take account of the smaller communication bandwidth of CC, the frequencies were transformed into percentages. This showed that CC had a much lower socio-emotional content than the other conference types. This applies to positive actions (12.2% versus 27.8%) as well negative actions (1.5% versus 6.0%). Conversely, CC showed a higher score in both types of task-related interactions, 'answers' (77.6% versus 61.4%) and 'questions' (8.7% versus 4.8%). A comparison of the percentages of IPA categories with the estimated norms of Bales (1970) showed that for CC positive as well as negative actions were much less frequent than the suggested norms while percentages for the superordinate category 'attempted answers' were higher than the norms. The figures of MM and FTF were within the norm or much closer to the boundaries of the range. Although Bales suggested that the norms should not be applied too strictly, he also believed that they would provide some form of evaluation of intragroup interaction (Bales 1970).

While for most categories there was little difference between FTF and MM, the MM group showed the highest level for socio-emotional behaviour. This observation is an exception to the overall pattern of

Table 2. Summary of experimental effects (FTF = Face-to-face; CC = Computer conference; MM = Multimedia; HSLT = Heidelberger Struktur-Lege-Technik).

Dependent variable	Effect of medium	Effect of elicitation method	Interaction
Conference duration	FTF shorter than two others	HSLT required more time than interviewing	Very long duration for MM employing HSLT
Quantity of generated knowledge	–	–	–
Quantity of generated concept relations	CC more productive than two others	HSLT more productive	–
Quality of knowledge generated	CC more effective than two others	HSLT more effective	MM shows poorest result for interviewing and best for HSLT
Quality of data for expert system	–	–	–
Interaction process analysis	Lower frequency of group interactions for CC than for two others; relative more task-related interaction and less socio-emotional content	More frequent group interactions under HSLT than for structured interviews	–

results, which consistently showed a strong similarity between FTF and MM for interaction data as well as performance indicators.

4. Discussion

The overall pattern of results suggested that there was only little difference between FTF and MM, compared to the CC condition. Not only does this apply to the performance data, it is also consistent with the results of the interaction process data. FTF and MM have in common that both use two communication channels (i.e. oral and visual) while CC is only based on written information. It seems that similarity in communication bandwidth is a more important factor than technical similarity of the communication process (i.e. no direct personal contact, computer-mediation), that is shared by CC and MM. As the present study provided a complex pattern of results, the most important findings are summarised in table 2.

This experimental work suggested that for the task selected, CC appeared to be the most appropriate medium since it came out best in a number of measures. As table 2 demonstrates, CC was most productive, indicated by quantitative as well as qualitative performance variables. This might be due to the high focus on facts and the low negative socio-emotional content of the discussions. The CC group predominantly gave information while MM and FTF showed a higher prevalence of giving opinions, as a more detailed IPA analysis revealed. Furthermore, due to reduced communication speed caused by keyboard input, contributions during the computer conference were rather short, which may have contributed towards focusing on facts rather than socio-emotional aspects of the discussion. All these factors may have strongly contributed towards good performance levels. Conversely, under MM and FTF conditions, the interaction was more person-oriented rather than being concerned with written statements, which are deprived of most non-verbal and para-verbal connotations.

A comparison of the IPA data with the estimated norms of Bales (1970) indicated that the results for MM and FTF corresponded to the norms (indicating typical communication patterns) much more closely than the data for CC. The lower socio-emotional content in CC may be partly due to the reduced number of transmission channels (e.g. non-verbal and para-verbal cues cannot be given). However, during CC where participants know each other well, socio-emotional aspects of the discussion may gain in importance (Walther and Burgoon 1992). Other work has also found between 20–30% of socio-emotional content out of total message content in CC (Rice and Love 1987). This suggests that outside the laboratory the difference between FTF and CC is likely to be smaller than found in this study.

At first sight the results seemed to indicate considerable advantages of CC over the other media. However, when examining performance measures at various stages of task completion, a somewhat modified picture emerged. It showed that the CC group did not manage to maintain their performance advantage when the final task (working out the design recommendations) was to be completed. As the difference index indicated, the CC group displayed a clear loss in performance. This raises the question of why it was not possible for the knowledge engineer to produce good design recommendations based on the good conference outcome. Whereas the computer conference outcome was based on a team effort, the design recommendations were based on the ability of the knowledge engineer to

transform the conference outcome into recommendations. This required knowledge on his part to complete this transformation. Based on the 'levels-of-processing' approach (Craik and Lockhart 1972, Craik 1973) and dual-trace theory (Paivio 1975), we would argue that multiple channel communication in FTF and MM may have allowed the knowledge engineers a deeper encoding of the information. This permitted them to close the performance gap to their colleagues using CC. While CC allowed the production and collection of more information, this may have been achieved only at the expense of a shallower processing of information.

MM was the most demanding condition since many different activities had to be completed simultaneously (reading text on screen, looking at graphics and video pictures, talking into the microphone, manipulating the keyboard) in quite an unfamiliar environment. This was reflected in the long completion times for MM under the HSLT condition, due to the combined demands of the more complex knowledge elicitation method and the most complex communication environment. In informal post-experimental interviews, participants reported that the effort required to integrate information provided by the MM system was considerable, making them feel very fatigued after the session. The multiple task environment might also have increased tension and arousal, indicated by more frequent negative socio-emotional behaviour found in the IPA analysis for MM sessions. Altogether, this was indicative of the additional demands imposed by the rather complex technical environment. Increased time requirements for using complex conference media is a typical finding in research into the effects of CMC (see Reid *et al.* 1996). Since negative effects of medium complexity will be reduced with increasing practice, organizations are advised to provide sufficient training and opportunities for practice concerning the use of modern communication media to achieve adequate performance standards.

The general pattern indicated that HSLT had some advantages over interviewing. Although the use of HSLT resulted in a longer conference duration, the conference outcome was generally better. While there was no difference between elicitation methods for the number of elicited concepts, the number of relations generated increased under HSLT. This suggested that HSLT was better at structuring the information but it did not generate more information. The distinction between quantity of information and quality of its structure resembles the typical difference found between the knowledge structures of a novice and an expert. While the latter encodes and structures the problem domain in a more effective manner, the total amount of knowledge may not be that different between the two (Eysenck and Keane 1995).

The judges' ratings of the conference outcome also showed better performance for the HSLT group. However, the advantage of HSLT disappeared when the final recommendations were to be made for expert system design. While interview-trained knowledge engineers showed significant improvements in transferring the good conference outcome into good design recommendations, the HSLT-group merely maintained their performance. This raises the important question of why this difference between elicitation techniques emerged. It seems that knowledge engineers using structured interviewing were more actively involved in the knowledge elicitation process because, in contrast to HSLT, interviews did not automatically give some structure to the collected data. The more active involvement may have resulted in a more elaborate encoding of the information (see Craik and Tulving 1975), which allowed interview-trained knowledge engineers to add significantly to the conference outcome. There was some degree of similarity between the result patterns for HSLT and CC. Both conditions supported the generation of knowledge but only at the cost of information processing depths, resulting in poorer performance for subsequent transfer activities.

Within the theoretical framework of Straus and McGrath (1994), the task used in the present study may be classified as an idea-generation task. In these kind of tasks, little coordination is required and little consensus needs to be reached. The theory suggests that CC is a very good medium for the type of task used in our experiment. However, as the results from our work have shown, multiple task requirements (i.e. running a productive conference session and deriving data for expert systems from the conference results) may increase the difficulty of selecting the best medium for a given task. While strong task orientation may be advantageous to solve certain problems, one needs to be aware that task-oriented conferences do not only reduce negative aspects of socio-emotional communication patterns but also show an impoverished positive socio-emotional interaction. This may be detrimental in situations, where positive group feedback is needed. For example, in the context of innovation management, the socio-emotional support and cohesiveness was judged to be an important factor for highly innovative groups (West 1990). Social context cues are also needed in situations where group members need to reach consensus which is based on their values and attitudes rather than on the identification of the 'best' solution (McGrath 1984, Straus and McGrath 1994). In product design teams, these social context cues may be needed during the process of 'finding common ground' (Steinheider *et al.* 1999), which is a critical activity to form a team with very diverse members (e.g. differences in

professional backgrounds, knowledge structures, approaches to problem-solving, language and attitude).

5. Implication

The main finding of this study has suggested that CC and HSLT were most effective in generating knowledge but this was only achieved at the cost of information processing depths, leading to poorer performance during a later transfer task. This and other findings have demonstrated the need for a very task-specific evaluation of communication media and knowledge elicitation methods.

This study and other work have also indicated that the effectiveness of communication media and elicitation methods has been conditional upon other variables. Organizations could use conference types more efficaciously by considering important contingency factors, such as task characteristics, goals of meeting, and so forth. In the context of product design, this suggests that CC may be most effective in those design activities that involve idea generation (e.g. during task clarification; see Pahl and Beitz 1996). In contrast, when judgements need to be made about which design proposal to adopt, conference types that provide stronger social context cues would be preferable. An example of this is the conceptual design phase (see Pahl and Beitz 1996), during which group members need to agree on a design solution after having discussed the pros and cons of each alternative.

Finally, with regard to the selection of the most effective elicitation technique, one would argue that a highly structured technique is more suitable if the knowledge engineer has little knowledge of the area and there is no benefit to enhancing it (e.g. novice only works on the project temporarily). Conversely, if it is critical that the knowledge base of the knowledge engineer is enhanced as a result of the elicitation process, structured interviewing appears to be more appropriate.

References

Bainbridge L. 1979, Verbal reports as evidence of the process operator's knowledge. *International Journal of Man-Machine Studies*, **11,** 411–436.

Bales, R. F. 1970, *Personality and Interpersonal Behavior* (New York: Holt, Rinehart and Winston).

Barua, A., Chellappa, R. and Whinston, A. B. 1997, Social computing: computer supported cooperative work and groupware, in G. Salvendy (ed.) *Handbook of human factors and ergonomics* (New York: Wiley and Sons), 1760–1782.

Cordingley, E. 1989, Knowledge elicitation techniques for knowledge-based systems, in D. Diaper (ed.) *Knowledge elicitation* (Chichester: Ellis Horwood), 87–175.

Craik, F. I .M. and Lockhart, R. S. 1972, Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, **11,** 671–684.

Craik, F. I .M. 1973, A 'levels of analyses' view of memory, in P. Pliner, L. Krames and T. M. Alloway (eds) *Communication and affect: Language and thought* (London: Academic Press).

Craik, F. I .M. and Tulving, E. 1975, Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, **104,** 268–294.

Eirlenspiel, K. 1995, *Integrierte Produktentwicklung. Methoden für Prozessorganization, Produktentwicklung und Konstruktion* (München: Hanser).

Eysenck, M. W. and Keane, M. T. 1995, *Cognitive Psychology Lawrence* (London: Erlbaum).

Finn, K. E., Sellen, A. J. and Wilbur, S. (eds) 1997, *Video-mediated communication* (Mahwah NJ: Lawrence Erlbaum).

Frankenberger, E. 1997, *Arbeitsteilige Produktentwicklung* (Düsseldorf: VDI-Verlag).

Hacker, W. and Jilge, S. 1993, Vergleich verschiedener Methoden zur Ermittlung von Handlungswissen. *Zeitschrift für Arbeits- und Organisationspsychologie*, **37,** 64–72.

Johansen, R. 1988, *Groupware: Computer support for business teams* (New York: Free Press).

Kirwan, B. and Ainsworth, L. K. 1992, *A Guide to Task Analysis* (London: Taylor and Francis).

Konradt U. 1992, *Analyse von Strategien bei der Störungsdiagnose in der flexibel automatisierten Fertigung* (Bochum: Brockmeyer).

McCarthy, J. C. and Monk A. F. 1994, Measuring the quality of computer-mediated communication. *Behaviour and Information Technology*, **13,** 311–319.

McGrath, J. E. 1984, *Groups: Interaction and performance* (Englewood Cliffs, NJ: Prentice Hall).

Pahl G. and Beitz, W. 1996, *Engineering design: a systematic approach* (London: Springer).

Paivio, A. 1975, Coding distinctions and repetition effects in memory, in G. H. Bower (ed.) *The psychology of learning and motivation*, Vol. 9 (New York: Academic Press).

Reid, F. J. M., Malinek, V., Stott, C. J. T. and Evans, J. ST. B. T. 1996, The messaging threshold in computer–mediated communication. *Ergonomics*, **39,** 1017–1037.

Rice R. E. and Love, G. (1987). Electronic emotion: socio-emotional content in a computer-mediated communication network. *Communication Research*, **14,** 85–108.

Rüttinger, B., Letter B. and Schramme, S. 1994, Computermediierte Kommunikation und Entscheidungsfindung, in R. Wille and M. Zickwolf (eds) *Begriffliche Wissensverarbeitung* (Mannheim: Wissenschaftsverlag), 289–300.

Scheele, B. and Gröben, N. 1979, *Die Heidelberger Struktur–Lege–Technik (SLT). Eine Dialog–Konsens–Methode zur Erhebung subjektiver Theorien mittlerer Reichweite* (Weinheim: Beltz).

Scott, A. C., Clayton, J. E. and Gibson, E. L. 1991, A practical guide to knowledge acquisition (New York: Addison-Wesley).

Sproull, L. and Kiesler, S. 1986, Reducing social context cues: electronic mail in organizational communication. *Management Science*, **32,** 1492–1512.

STEINHEIDER, B., BURGER, E., GANZ, W. and WARSCHAT, J. 1999, A model to support expert cooperation. Paper presented at 32nd ISATA, International Symposium on Automotive Technology and Automation, 14–18 June 1999, Vienna, Austria.

STRAUS, S. G. 1996, Getting a clue: the effects of communication media and information distribution on participation and performance in computer-mediated and face-to-face groups. *Small Group Research*, **27,** 115–142.

STRAUS S.G. and MCGRATH, J. E. 1994, Does the medium matter? The interaction of task type and technology on group performance and member reactions. *Journal of Applied Psychology*, **79,** 87–97.

WALTHER, J. B. and BURGOON, J. K. 1992, Relational communication in computer-mediated interaction. *Human Communication Research*, **19,** 50–88.

WALTHER, J. B. 1992, Interpersonal effects in computer-mediated interaction: a relational perspective. *Communication Research*, **19,** 52–90.

WALTHER J. B. 1995, Relational aspects of computer-mediated communication: experimental observations over time. *Organization Science*, **6,** 186–203.

WEISBAND, S. P., SCHNEIDER, S. K. and CONNOLLY, T. 1995, Computer-mediated communication and social information: status salience and status differences. *Academy of Management Journal*, **38,** 1124–1151.

WEST M. A. 1990, The social psychology of innovation in groups, in M. A. West and J.L. Farr (eds) *Innovation and Creativity at Work* (London: Wiley).