Experimental Comparison of Multimodal Meeting Browsers

Wilfried Post, Erwin Elling, Anita Cremers, and Wessel Kraaij

TNO Human Factors, Kampweg 5, 3769 DE Soesterberg, The Netherlands
{Wilfried.Post, Erwin.Elling, Anita.Cremers, Wessel.Kraaij}@tno.nl

Abstract. This paper describes an experimental comparison of three variants of a meeting browser. This browser incorporates innovative, multimodal technologies to enable storage and smart retrieval of captured meeting. Over a hundred subjects had to work in a design team in which they had to prepare and carry out a final meeting, while supported by one of the browser variants. In one condition, teams worked without such support. Measures on individual characteristics, the team, the process and outcome of the project, and the usability of the browsers were taken. The results indicate that a multimodal meeting browser can indeed improve meetings. Further analysis of the now available data will provide additional insight into how browsers can contribute to more efficient and satisfactory meetings, improved team performance and higher quality project outcomes.

1 Introduction

The research reported here is carried out as part of AMI, a European Integrated Project that aims at developing new multimodal technologies for supporting human interaction in the context of meetings [1]. These technologies enable storage, interpretation and retrieval of captured meeting interactions and will be used to develop a meeting browser that should make meetings more effective, efficient and pleasurable. The importance of improving meetings is evident when is realized that most professionals join (but not enjoy) at least one meeting a day [2]. Not so evident, however, is to establish that a meeting browser indeed improves meetings, and on which elements this depends. The objective of the presented study is to determine whether and how a multimodal meeting browser improves a meeting. Our ultimate aim is a more efficient and satisfactory project process and higher quality results.

2 Method

An experiment was set up to compare meetings with and without meeting browser support. For a better understanding of what form of support might work best, in total three variants of a multimodal meeting browser were part of the comparison. The setup is described next.

2.1 Stimulus material

In [3] is pointed out that the success of a meeting is better determined from a series of meetings, such as in a project with a clear goal. Further, the success of a meeting, or a project, depends not only on the means used (e.g., a meeting browser), but also on (project or meeting) method, individual factors, team factors, type of task, organizational culture, environment, etc. These factors have been specified in the following experimental scenario. The subjects have to participate in a design project team, playing a specific role (project manager, industrial designer, user interface designer and marketing expert). They are told to take over a project carried out so far by a team that did not do well enough. The subjects have to use all the material used and produced by this previous team. They have to prepare and carry out a final design meeting in which they have to design a television remote control, according to specific requirements. Both preparation and execution of the meeting is carried out in meeting rooms in Soesterberg (Fig. 1) and Edinburgh, well instrumented research environments for four subjects, with individual workplaces (including a private computer), a shared workplace (including electronic presentation boards), and, depending on the experimental condition, a particular set of tools. The participants and their computer interactions are observed and recorded by means of video cameras, microphones, and screen videos.

On preplanned points in time, subjects receive e-mails about the tasks to carry out (sent by a virtual head of the department), some hints (sent by a virtual coach), and a series of questionnaires and rating scales.

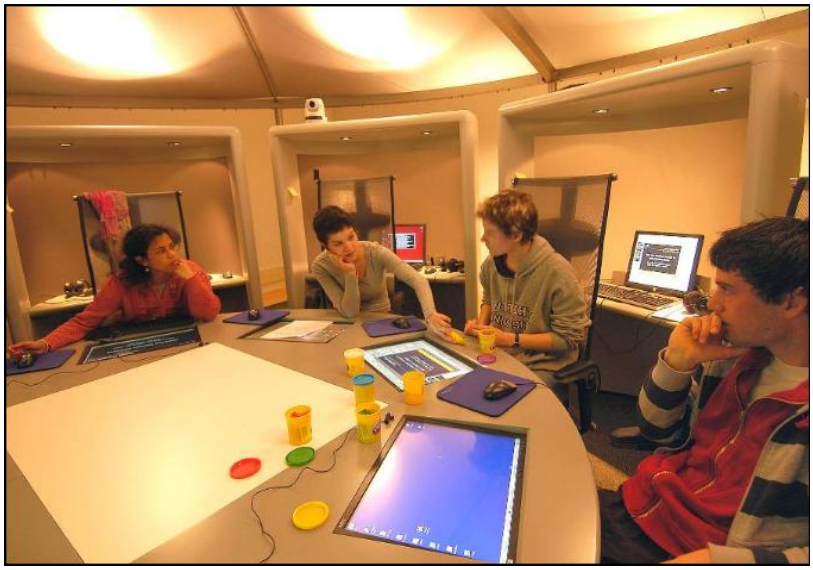

Fig. 1. Research environment called “Team Cockpit”. On the left, not shown, is a shared large presentation screen. The subjects are discussing a clay prototype remote control.

2.2 Subjects

Data from 27 project teams consisting of 4 participants were collected, of which 12 in Soesterberg and 15 in Edinburgh. Most of the 108 subjects were undergraduate students. All of them were paid €45,- for about 4 hours of work.

2.3 Independent variable

The four experimental conditions are as follows.

1. In the basic condition, no computer support is provided, other than a folder structure in Microsoft Explorer, organized by project phase (Project Kick-off; Functional Design Phase and Conceptual Design Phase). In these folders, they can find documents, minutes, PowerPoint slides and audio/video recordings of three previous meetings.
2. In the second condition, all documents are still provided in a folder structure. For making use of the meeting registrations, however, a meeting browser is provided, which offers synchronization of the multimodal meeting recordings (i.e., synchronous browsing through PowerPoint slides, automatic speech transcripts or audio/video material).
3. The third condition is similar to the second one, except that the browser includes automatic generated abstracts - new technology that will be available in the near future. Fig. 2 shows a screenshot of the browser. In the upper middle part of the screen, one can browse through audio/video registration of a particular meeting within a project. In the lower left corner, one can browse through speaker segmentations: each participant has its own speaker activity log represented. In the lower right part, an automatic transcript is presented, which is browsable by scrolling and searchable by keyword search. In the upper left corner, the PowerPoint slide of a particular episode that is browsed is shown. All these parts are synchronized.
4. In the fourth and final condition, all material, including documents, are integrated from the perspective of a user in a task setting (i.e., carrying out role-specific work in a design project (see [4]). The Task Based Project Browser, as it is called, provides direct access to three different information sources via the tabs Meetings, Documents and Messages. In addition, it is possible to access these sources indirectly via three task-oriented tabs: *Project* (project details, people involved, different design phases), *Todo's* (see Fig. 3) and *Decisions*. All information items are hyperlinked to the original sources. This makes it possible to, for instance, immediately view a meeting clip in which a specific information item, such as an action item, is being discussed. The information provided in the Meetings tab is identical to the information in the meeting browser of the third condition.

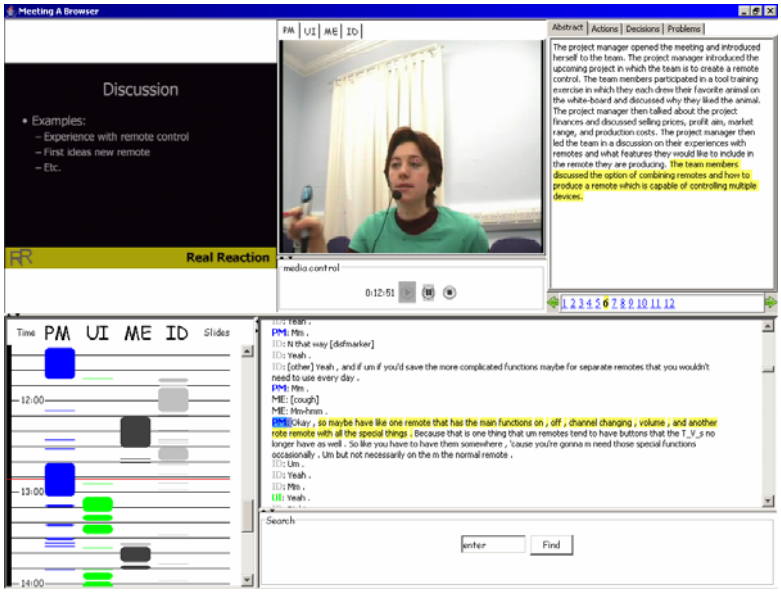

Fig. 2. Screenshot of the meeting browser used in condition 3. The browser used in condition 2 is similar, except for the automatic generated abstracts in the upper right corner.

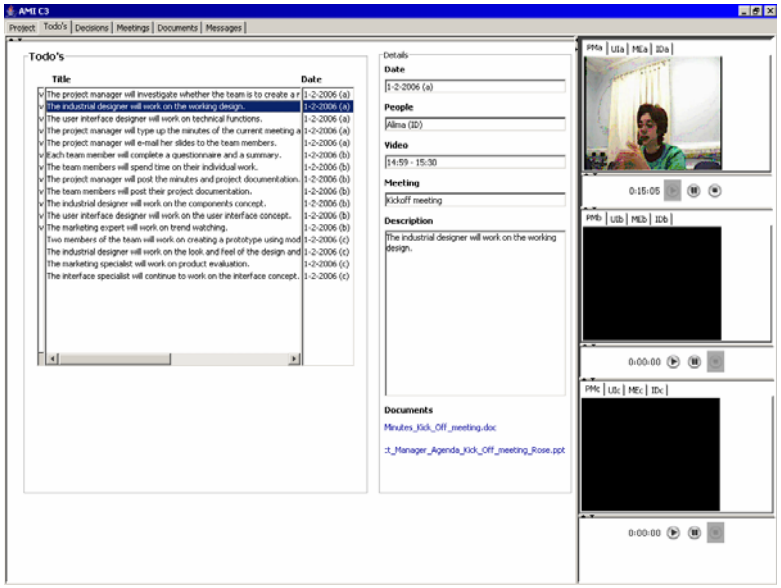

Fig. 3. Screenshot of the “Todo’s” tab of the Task Based Project Browser used in condition 4.

Table 1. Overview of how in the four experimental conditions the various sources of information can be retrieved with the four types of tools: the system, Email, Internet browser, and Multimodal Meeting browser.

Information	**Condition 1**				**Condition 2**				**Condition 3**				**Condition 4**			
	F	E	I	B	F	E	I	B	F	E	I	B	F	E	I	B
Minutes	x				x				x			x				x
Presentations	x				x			x	x			x				x
Emails/messages		x				x				x						x
Internet			x				x				x					x
Close up videos	x							x				x				x
Speaker activity log								x				x				x
Meeting transcripts								x				x				x
Abstracts	x				x				x			x				x
Actions/Todo's	x				x				x			x				x
Decisions	x				x				x			x				x
Problems/Open issues	x				x				x			x				x

2.4 Procedure

After the subjects were welcomed, the experiment was explained. Next, they were asked to divide the four team roles. They all had to open their e-mail browser and from that point on, they had to follow the instructions sent to them. To prevent bias caused by lack of familiarity or skill with the tool set, the subjects were trained first. This started with reading a general explanation of the functionalities of the tool set the subjects were assigned to. Next, the subjects received e-mail instructions to answer specific questions for which they had to exercise with various parts of the tool set. The training lasted about 30 minutes.

The experimental part consisted of three phases. The first two phases were carried out individually. First, they had to familiarize themselves with the project, the previous team and their personal roles, which will be further referred to as "getting a gist". They had to do this in 15 minutes. Second, they had to prepare the upcoming meeting in 45 minutes. Finally, they had to join the meeting, which lasted 45 minutes. Before, in between and after the phases, the subjects received e-mails that were either links to electronic questionnaires or task related instructions. An overview of the questionnaires is found in Table 2. During the second and third phase, they received a warning five minutes before the end of it.

2.5 Dependent variables

A specially developed evaluation instrument was used for measuring project process and outcome (see [5]). This instrument includes subjective workload rating scales and team questionnaires, and objective analysis of information transfer and project outcome. An overview is provided in Table 2.

Table 2. Overview of all the measurements taken during the experiment

When	What	How
Individual characteristics		
Before the experiment	Background, experience	Questionnaire
	Spatial orientation	Test
	Memory	Test
Team and project measures		
During the meeting	Behavior	Subjective observation
After the meeting	Dominance	7 pt scale
	Info processing	4 items
	Leadership	4 items
	Process satisfaction	3 items
	Cohesiveness	5 items
	Work pace	4 items
	Communication	4 items
	Supporting behavior	8 items
	Effectiveness	4 items
	Efficiency	7 items
	Outcome satisfaction	5 items
	Team satisfaction	2 items
	Project experience	3 questions
Analysis afterwards	Info transfer	# shared info
	Info outcome	# correctly applied info
	Quality product	Multiple Expert Assessment
Usability measures		
Before getting a gist	Mental effort	150 pt scale
During getting a gist	Use of browser	Logging
	Behavior	Subjective observation
After getting a gist	Browser usability	Questionnaire
	Mental effort	150 pt scale
During preparation	Use of browser	Logging
	Behavior	Subjective observation
After preparation	Browser usability	Questionnaire
	Mental effort	150 pt scale
During meeting	Use of browser	Logging
	Behavior	Subjective observation
After the meeting	Browser usability	Questionnaire
	Mental effort	150 pt scale
	Usability browser	Questionnaire
	Meeting room of the future	Questionnaire

3 Results

3.1 Individual characteristics

In total, 108 subjects participated in the experiment. Eight teams of four subjects each were assigned to condition 1, seven to condition 2, seven to condition 3, and five to condition 4. Their mean age was 23 years old (s.d. 7 years), 56% were male and 44% female. Almost all participants were students (97%). Most of them were students of computer science or information science (50%), 16% were students of psychology, 10% of philosophy and the rest of varying or unspecified subjects.

Participants were asked a number of questions on their patterns of computer use. All participants use both the computer and the internet on a daily basis. Participants also received questions on their experience with meetings. Half of the participants participate in meetings on a weekly basis, one third on a monthly basis, one sixth never. More than half of the participants feel that most of the times objectives for their meetings are attained; one third feels that they are sometimes attained; one-sixth varies between the rest of the answers.

About 80% of the participants feel that either sometimes or most of the time, time for their meetings is well-spent. About half of the participants feel that most of the times they like to participate in meetings. The others either like it sometimes or always. Hardly anybody never likes to participate. When asked to characterize their typical meetings, about 80% answer that the meetings have an informal atmosphere, about 10% formal. Other qualifications used are either positive (good: 6%) or negative (tedious: 4%).

Finally, participants were asked about their experience with working in project teams. About 80% has either hardly any or average experience with working in project teams. The rest has either no experience or a lot of experience. About half of the participants have no experience at all in product or service development. The other half has hardly any or average experience.

3.2 Team and project measures

For the variables *Information processing*, *Leadership*, *Process satisfaction* and *Outcome satisfaction*, that were scored on a 7 point scale, notable results are found. *Information processing*, *Leadership* and *Process satisfaction* are in condition 3, significantly higher than in condition 1 (resp. $F(3,26)=3.621$; $p<5$; , $F(3,26)=3.015$, $p=0.5$); $F(3,26)=3.253$, $p=0.5$), and *Outcome satisfaction* is almost significantly higher ($F(3.26)=2.843$, $p=0.6$).

3.3 Usability

At three points in time during the scenario participants were asked fill out the browser usability questionnaire with questions about their opinion on the tool set that were available to them. Questions were asked referring to the usability aspects effectiveness, efficiency and satisfaction. For every aspect four questions were asked, in a random order. Participants answered the questions on a seven-point scale, varying from not applicable at all (1) to very much applicable (7).

Mean total scores per assessment and per condition were calculated. All mean scores varied between 3.45 and 4.67, indicating general moderate usability. There were no differences between scores for effectiveness, efficiency and satisfaction.

Significant effects of point of measurement and condition occurred. For all three aspects there was a significant effect of condition. For effectiveness, condition 1 differed significantly from both condition 2 and 3 ($F(3,23)=6.39$; $p<.01$). For efficiency, condition 1 differed significantly from condition 3 ($F(3,23)=3.68$; $p<.05$). For satisfaction, condition 1 differed significantly from 3 (and almost from 4) ($F(3,23)=5.17$; $p<.01$).

For all three aspects a significant effect of point of measurement occurred. For effectiveness, measurement point 1 differed significantly different from measurement point 3 ($F(2,46)=6.39$; $p<.01$). For efficiency, measurement point 1 differed significantly different from measurement point 3 ($F(2,46)=4.37$; $p<.05$). For satisfaction, measurement point 1 differed significantly from both measurement points 2 and 3 ($F(2,46)=11.04$; $p<.001$). This can be considered a learning effect.

Table 3. Mean effectiveness, efficiency and satisfaction for condition 1 to 4.

Effectiveness				**Efficiency**				**Satisfaction**			
1	2	3	4	1	2	3	4	1	2	3	4
3.45	4.26	4.62	4.19	3.67	4.07	4.40	4.17	3.74	4.18	4.67	4.47

Table 4. Mean effectiveness, efficiency and satisfaction after familiarization (1), preparation (2) and meeting (3).

Effectiveness			**Efficiency**			**Satisfaction**		
1	2	3	1	2	3	1	2	3
3.95	4.16	4.27	3.89	4.11	4.22	4.02	4.27	4.50

3.4 Afterwards

Afterwards, participants were asked three questions on their experiences working in the project, giving them five possible answers: never, hardly ever, sometimes, most of the times, and always. Answers were generally positive. The mean score over all conditions on the question whether they felt the objectives for the project had been attained, was 3.82 (most of the times). There were no significant differences between the conditions. The mean score over all conditions on the question whether they felt

the time for the project had been well-spent was 3.74 (most of the times). There were no significant differences between the conditions, although the difference between conditions 1 and 3 was almost significant ($F(3,23)=3.00;p=.054$). The mean score over all conditions on the question whether they had generally liked to participate in the meeting was 4.00 (most of the times). There was a significant difference between the conditions 2 and 3 $F(3,23)=4.11;p<.05$). The difference between conditions 2 and 4 was almost significant ($p=.057$).

3.5 Subjective observations

Up to now we have only presented objective evidence, but some subjective material has been collected as well. Some personal impressions of the experimenter:

- In all conditions, there was a clear learning curve. Participants were mostly hesitant at the start. This gradually decreased and led to curiosity towards going straight to the information they were looking for.
- In all conditions, a small number of people suffered from information overload. They did not know where to start looking for, or lost track. It should be mentioned that the design task is quite complex.
- The browser variants were mainly used during *getting a gist* and *meeting preparation*, and not during a *meeting*.
- Whatever condition, it appeared that without a strong, goal-oriented, natural leader, the team process was slower and fewer ideas and information were exchanged.
- Only in the task based meeting browser condition, some groups seemed to have time left for monitoring meeting progress. Also, only in this condition, the norm solution was found: a remote control meeting all criteria, including the form of a banana.

6 Conclusions and future work

The most important result of this work is that we have found experimental evidence that a meeting browser indeed can improve a meeting. Especially meetings supported by the browser used in condition 3 have a significant higher perceived effectiveness, efficiency and satisfaction. Our expectation, however, that meetings supported with the Task Based Project Browser would be superior could not be shown. We provide some possible explanations for this result.

During experimentation, we were conscious of the limitations of the current experimental implementation of the Task Based Meeting Browser. Due to time and budget constraints, it contains less functionality and is less integrated than we originally had conceived (see [4] for a full description). Subjects regularly mentioned additional functionalities, alternative ways for organizing all information, other search

possibilities - suggestions that were already part of our conceptual design. These are taken into account during further development.

Working on a design project is difficult, especially for our inexperienced and rather young subjects. This may have influenced the results in the sense that they were probably not fully aware of the requirements for meetings and functioning in a project team. They were experienced with using the computer, though, which suggests potential skills for using the tools and trust in the information offered. This could be a reason why they were quite apt at using the tools and were not really intimidated by the prospect of meeting in an instrumented meeting room.

A huge amount of data is currently waiting for analysis, such as the influence of individual characteristics, information flow analysis, using expert opinions about the resulting clay remote control prototype as outcome measure, and detailed analysis of tool usage using our logging data. We hope to see more beneficiary effect of the meeting browser, and expect to find some differences between the three variants. This would provide us explanatory material, and subsequently guidance for further development. Interesting is the effect we have found of experimental condition on *leadership*. If this effect is strongly related to the outcome of a project, this would indicate that a meeting support tool should focus more on, for example, "leadership support". Multimodal technologies can help recognize various types of social interactions, and can accordingly be applied to support social behavior itself.

Acknowledgments. This work was partly supported by the European Union 6th FWP IST Integrated Project AMI (Augmented Multi-party Interaction, FP6-506811, publication AMI-234). About half of the data collection has been carried out at the University of Edinburgh, thanks to Jean Carletta, Mike Lincoln and Melissa Kronenthal. Charley Penning, Erik Boertjes, Jonathan Kilgour, Mike Flynn and Stephan Lesch contributed substantially to development of the meeting browser variants.

References

1. AMI (2003). Augmented Multi-party Interaction (AMI). Sixth Framework Programme, Priority 2, Information Society Technologies, Integrated project Annex 1 – "Description of Work".
2. Rogelberg, S. G., Leach, D. J., Warr, P. B. & Burnfield, J. L. (in press). "Not another meeting!" Are meeting time demands related to employee well-being? Journal of Applied Psychology.
3. Post, W.M., Cremers, A,H.M. and Blanson Henkemans, O. (2004). A research environment for meeting behavior. In: A. Nijholt & T. Nishida (Eds.). Proceedings of Social Intelligence Design 2004, p. 159-165.
4. Cremers, A.H.M., Groenewegen, P., Kuijper, I & Post, W.M. (2007). The Project Browser: supporting information access for a project team. This volume.
5. Post, W.M., Huis In't Veld, M.A.A. and van den Boogaard, S.A.A. (2008). Evaluation Meeting Support Tools. In: Personal and Ubiquitous Computing. Accepted for publication.