An Empirical Investigation on Text-Based Communication in Distributed Requirements Workshops

Fabio Calefato
Dipartimento di Informatica
University of Bari, Italy
calefato@di.uniba.it

Daniela Damian
Dept of Computer Science
University of Victoria, Canada
danielad@cs.uvic.ca

Filippo Lanubile
Dipartimento di Informatica
University of Bari, Italy
lanubile@di.uniba.it

Abstract

Among the software development activities, requirements engineering is one of the most communication-intensive and then, its effectiveness is greatly constrained by the geographical distance between stakeholders. For this reason, the need to identify the appropriate task/technology fits to support teams of geographically dispersed stakeholders plays a key role for coping with the lack of physical proximity when developing requirements.

In this paper we report on an empirical study that assessed the use of synchronous text-based communication in distributed requirements workshops, as compared to face-to-face (F2F), and the effects of computer-mediated communication (CMC), with respects to the different tasks of distributed requirements elicitation and negotiation. First results show that, in terms of satisfaction with performance, CMC elicitation is a better task/technology fit than CMC negotiation. Furthermore, the general preference for F2F over CMC is due to the strong preference for the F2F negotiation fit over the CMC counterpart.

1. Introduction

Over the past three decades, and particularly in the mid-1990s, many experimental studies on deployments of both desktop and classroom videoconferencing have been published. Some of these studies report about successful interaction among the remote sites, with no losses compared to face-to-face (F2F) interaction [24],[18], whereas others describe failures due to technical and behavioral issues [27],[16],[23]. Today, despite the recent advances in video and audio technology and the increasing ability to create a rich medium for distributed meetings, the practicality of organizing videoconferences still remains low, due to the considerable overhead. The necessary infrastructure is expensive, awkward to setup and maintain at remote sites, and its coordination across organizational boundaries is often problematic [25].

While there is an interesting body of knowledge about the comparison between F2F and audio/video technology, although with mixed results, past research on media effects has not given the same attention to the comparison between F2F and synchronous, text-based interaction. Such disregard is probably due to the many theories of computer-mediated communication (CMC) [9], which recommended the use of rich media for complex tasks as the only possible solution. However, prominent theories such as Media Richness [7] and Social Presence [26] have strong face validity, but empirical evidence is rather equivocal [10]. A number of studies of media use have provided evidence that runs counter to the predictions, particularly when media other than F2F communication are utilized, thus pushing researchers to theorize that media selection is also affected by factors beyond richness [4]. Such theories have fallen short when considering context and task complexity for media selection. The existing literature on Group Support System (GSS, see [12] for an exhaustive compendium) has often reported of distributed groups who, while interacting via text-chat, outperformed collocated groups in idea generation tasks, but were outperformed in problem-solving tasks [21]. More recently, Birnholtz et al. proved the existence of collaboration settings, characterized by reduced information loads, where synchronous, text-based communication was adequate to achieve common ground among conversational participants unknown to each other [1]. These results suggest that CMC theories cannot be accepted or considered valid *tout court* and that an analysis of the appropriateness of the fit between task characteristics (e.g., complexity) and technology characteristics (e.g., medium synchronicity

International Conference on Global Software Engineering(ICGSE 2007)
0-7695-2920-8/07 $25.00 © 2007 IEEE

and richness levels) is needed to get the best out of media use [30]. Further, a common limitation of CMC empirical studies is the evaluation of media effects on the execution of generic tasks, whereas executing realistic tasks requires individuals to apply known techniques or recall specialized knowledge to be performed [21].

The goal of the empirical investigation described in this paper is to evaluate (1) the use of synchronous, text-based communication in distributed requirements workshops, as compared to F2F, and (2) the effects of CMC with respects to the different tasks of distributed requirements elicitation and negotiation.

Requirements engineering is an appropriate domain for this study for a couple of reasons. First, it involves a complex set of communication-intensive tasks. Requirements elicitations and negotiations are among the most challenging and communication-intensive practice in software engineering [19]. Further, requirements elicitation and negotiation are complex tasks that require a constant interplay between idea generation, decision making, and conflict resolution activities, although in different measure (elicitation is more a generative task, whereas negotiation is more oriented to decision making). Secondly, recent research in the field has compared to F2F both audio and video links [17],[8], but it has not yet given same attention to synchronous, text-based communication.

The paper is organized as follows. Section 2 de describes the experiment in detail, including the design, instrumentation, data collection, measures and execution. Section 3 presents the results from data analysis. Section 4 discusses the findings from the experiments, whereas Section 5 discusses the threats to validity. Finally, conclusions are presented in Section 6.

2. The Experiment

We conducted an empirical study of six academic groups, playing the role of stakeholders involved in requirements engineering activities. The six groups observed (Gr1-6) were attending a Requirements Engineering course held at the University of Victoria in 2006. The study subjects were forty undergraduate students who volunteered to take part in the experimentation, after giving informed consent. Each group was composed of five to eight randomly-selected students (the terms students, stakeholders, and study participants are used interchangeably henceforth). Furthermore, the projects were randomly assigned to groups before group membership was determined. Each of the six software projects was developed through the interaction of a client and a developer team. Table 1 shows the student groups assigned to the six project teams. As an educational constraint imposed by the course, the project assignment was done so that each student was involved in two projects at the same time, as either client or developer. For instance, students belonging to Gr1 acted as clients in Project1, and as developers in Project6.

Table 1. Groups and allocation to projects

Project	Client team	Developer team
Project1	Gr1	Gr2
Project2	Gr2	Gr3
Project3	Gr3	Gr4
Project4	Gr4	Gr5
Project5	Gr5	Gr6
Project6	Gr6	Gr1

The goal of each project team was to develop a Requirements Specification (RS) document as a negotiated software contract between the developers team and the client team. The project work did not contemplate the writing of any code for the developer groups. Figure 1 illustrates the workflow of the requirements development process, over a period of about ten weeks. It comprises ten phases of continuous requirements discovery and validation, through which the understanding and documentation of requirements was improved. Each of these phases consists of tasks for either one of the client/developer groups, or both groups (project tasks). The developers, together with the clients, created several versions of the Requirements Specification document, while applying techniques of requirements elicitation and negotiation.

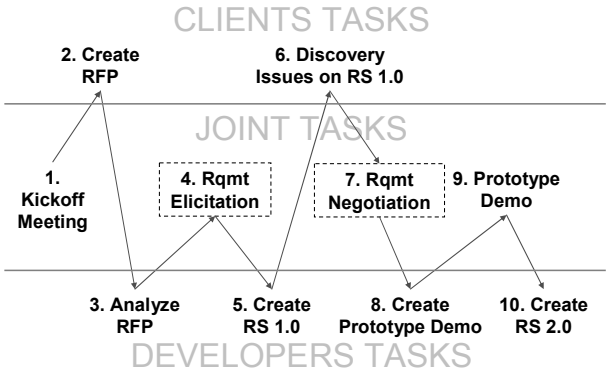

Figure 1. Workflow for the development process of the RS documents

The deliverables on which students were graded in the course are the RS 1.0 and 2.0, reflecting the shared

understanding of the project that the clients and the developers built over the requirements elicitation and negotiation workshops.

2.1. Design

The experiment requires comparing CMC and F2F communication mode in requirements elicitation and negotiations workshops. Table 2 shows the experimental plan, which corresponds to a 2^3 factorial design [20]. The three factors, each having two levels, are:

1. **communication mode** (levels: *F2F* and *CMC*);
2. **requirements workshop** (levels: *elicitation* and *negotiation*);
3. **role** (levels: *client* and *developer*).

The stakeholder-related observations, shown in groups for better readability, are the unit of analysis for this empirical design.

Table 2. The 2^3 factorial design of the experiment

	A **Comm Mode**	*B* **Rqmt Workshop**	*C* **Role**	**Subjects**
(1)	F2F	elicit	client	Gr1, Gr3, Gr5
a	CMC	elicit	client	Gr2, Gr4, Gr6
b	F2F	negot	client	Gr2, Gr4, Gr6
ab	CMC	negot	client	Gr1, Gr3, Gr5
c	F2F	elicit	dev	Gr2, Gr4, Gr6
ac	CMC	elicit	dev	Gr1, Gr3, Gr5
bc	F2F	negot	dev	Gr1, Gr3, Gr5
abc	CMC	negot	dev	Gr2, Gr4, Gr6

In the experiment, the *communication mode* and *requirements workshop* factors vary within subjects, whereas *role* factor varies between subjects. For instance, subjects in *Gr1* interacted as *clients* in *F2F elicitation* workshop (treatment combination *(1)*), and in *CMC negotiation* workshop (treatment combination *ab*). Conversely, they participated in *CMC elicitation* and *F2F negotiation* as *developers* (treatment combinations *ac* and *bc*, respectively). Albeit in different roles, with this experimental design we obtained data from the subjects for comparing CMC to F2F communication for the purpose of conducting requirements elicitations, as well as negotiations.

2.2. Instrumentation, Training, and Execution

The requirements workshop sessions were instructed so that all the workshops could be held in parallel and be completed within an hour. F2F workshops (both elicitations and negotiations) were held in parallel, in the same classroom. Also the CMC workshops were all held in parallel, but the students interacted from three different laboratories, so as to simulate geographical dispersion. Each student was assigned to a given seat, so that to avoid whole teams to stay in the same laboratory, and some participants in the same workshop to sit side by side. Due to course constraints, F2F and CMC requirements elicitation sessions involved two developers and the whole client team, whereas F2F and CMC negotiations involved the whole project teams (i.e., all the clients and developers).

CMC workshops were run using the eConference tool, a text-based, distributed meeting system [3]. To let participants gain familiarity with the tool, a one hour demo was given at class time. In addition, a user manual was made publicly available on the course web site. Furthermore, to reduce the risks of technical problems, a training session was instructed one week before each CMC workshop session, during which the students installed the tool and got acquainted with it.

During the execution of the CMC workshops, one of the researchers, a teaching assistant, and a Ph.D. student stayed in each laboratory to provide technical support, and to ensure that no participant verbally interacted with the others. It was fundamental to the study that the participants of the CMC sessions did not have access to any visual or verbal cues, unavailable in text-based communication. Furthermore, since the tool also supports IM, we decided to disable the roster management, so that the students were not able to add buddies to the contact list and chat "off topic" with their friends during the workshops.

2.3. Data Collection

The data sources for the experiment are the post-elicitation and post-negotiation questionnaires, which were administered to the students about one week after each requirements workshop session. The students received the two post-hoc questionnaires in both electronic and printed form. Students who returned the post-elicitation questionnaire were 20 out of 24 participants in total (83%), whereas the response rate for the post-negotiation questionnaire was lower (19 out of 38, 50%).

The questionnaires were formulated taking into account the communication issues commonly experienced and already acknowledged by previous research in the requirements engineering field [1], and the issues informally reported by the students after each requirements workshop session.

2.4. Dependent Variables and Measures

Satisfaction questionnaires are the only data source of the investigation considered in this report. Subjects' responses were then, coded to perform quantitative analysis.

To evaluate the differences between the requirements workshops and the communication modes through the subjects' perception, we conceptualized two constructs, namely (1) *satisfaction with performance* and (2) *comfort with communication mode*, adapted from [21].

With regard to the construct of satisfaction with performance, we defined a first 4-point Likert scale, anchored with '4=strongly agree,' and '1=strongly disagree.' The scale items aimed at weighing subjects' perception of the extent to which the decisions were consensus based and the amount of information generated was properly processed. We chose these two criteria because idea generation and consensus attainment are the dominant activities executed, respectively, when performing the tasks of eliciting and negotiating software requirements. The subjects provided responses to the each question in the scale for both F2F and CMC.

With respect to comfort with communication mode, we defined a second 5-item, 4-point Likert scale that aimed at assessing the perceived degree of discussion contentment and engagement level. We selected these criteria because we wanted to assess how media affect the opportunity to actively participate in the discussion and openly discuss conflictual issues.

To ensure the validity of the constructs, principal component analysis was applied. Principal component (or factors) analysis is a procedure that discards poorly-correlated questions and retains only those that account for a large amount of the total variance in the components data set, thus confirming the existence of the hypothesized components [13]. We also performed scale reliability analysis to further determine the internal construct validity by assessing the extent to which a set of questions measures a single latent variable. We used the Cronbach's alpha coefficient, the most-widely used index of internal consistency in social sciences [6].

3. Results

We report the results from the analyses applied to data collected from the subjects who got exposure to all the four workshop/medium combinations. We applied nonparametric statistics because the sample was rather small and we could not rely on the normality assumption. With respect to the construct of satisfaction with performance, we executed the Friedman test on the response set of the first scale, as a non-parametric alternative to the within-subjects analysis of variance for multiple dependent samples [5]. The purpose of applying this statistic is to determine whether there are significant differences in the level of subjects' satisfaction with performance between the four task/technology (i.e., workshop/medium) fits. In this analysis, the role factor is confounded with the interaction between the communication mode and requirement workshop factors. For each subject, first the responses were summed so as to obtain an overall score of the personal level of satisfaction with performance during the requirements workshops. Then, the ranks of the four workshop/medium fits were calculated on each per-subject summed scores (4^{th} rank corresponds to the highest score, 1^{st} rank to the lowest). The box plot in Figure 2 shows F2F negotiation to exhibit the highest, or best, mean rank (3.5) followed by F2F elicitation (2.75). CMC elicitation and CMC negotiation have the lowest average ranks (2.15 and 1.6, respectively). In addition, F2F and CMC negotiations exhibit a smaller rank variability compared, respectively, to F2F and CMC elicitations. The null hypothesis for the Friedman test is that the distribution of the ranks for each combination is the same. The test result indicates a statistically significant difference between the ranks at the 5% significance level (χ^2=14.54, p=.002) and, consequently, the null hypothesis is rejected.

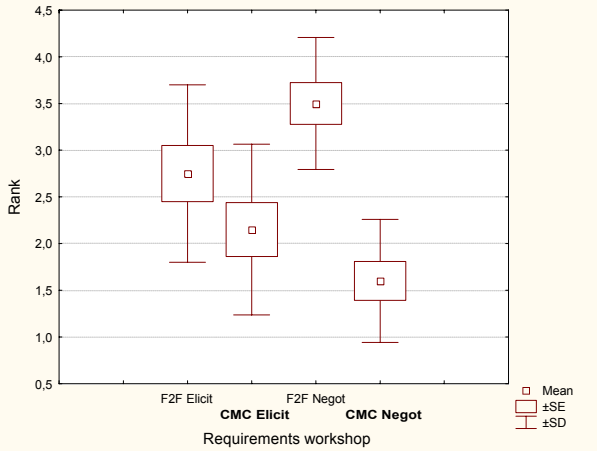

Figure 2. Ranks based on subjects' evaluation of satisfaction with performance (the higher the rank, the better the workshop/medium fit)

To further assess the differences between the ranks of the four workshop/medium fits, we applied a series of statistics to these scores to perform matched-pair

comparisons between (I) F2F elicitation and F2F negotiation, (II) F2F elicitation and CMC elicitation, (III) F2F negotiation and CMC negotiation, and, finally, (IV) CMC elicitation and CMC negotiation. The comparisons were performed by applying the Wilcoxon signed-rank test, as a nonparametric alternative to the t-test for two dependent samples [5]. The results, shown in Table 3, report for each matched-pair comparison (e.g. F2F elicitation vs. CMC elicitation), positive ranks (e.g., how many subjects preferred F2F excitation over the CMC counterpart), negative ranks (e.g., how many subjects preferred CMC elicitation over the F2F counterpart), and ties (e.g., how many subjects perceived F2F and CMC workshops to be equal). The Wilcoxon test for the first pair (I) resulted significant at the 5% level ($Z=2.27$, $p=.023$), showing a significant preference of subjects for F2F negotiations over F2F elicitations. The second and third Wilcoxon tests show that, while subjects significantly prefer F2F negotiation over CMC negotiation (III, $Z=2.54$, $p=.011$), no statistically significant difference was found in the comparison between F2F elicitation and CMC elicitation (II, $Z=1.56$, $p=.119$). Finally, the comparison between CMC elicitation and CMC negotiation was not found statistically significant as well.

Given the results of Wilcoxon test and Friedman test, we can conclude that study subjects perceived F2F negotiations as the best-fitting task/technology match in terms of the extent to which discussion was consensus-based and the information generated not missed.

With regard to the construct of comfort with communication mode, we applied principal component analysis to the second Likert scale defined in both the post-elicitation questionnaire and post-negotiation questionnaire. The analysis, performed with varimax rotation and a cut-off point of .70, extracted two identical components, retaining the same three items. The Cronbach's alpha index computed was .82 for the component extracted from the scale in the post-elicitation questionnaire, and .75 for the component extracted from the scale in the post-negotiation questionnaire. Both indexes are above the threshold of .70 suggested by Nunnally to affirm scale reliability [22].

Table 4 shows the breakdown of the responses to the items in the components extracted and the results of the chi-square goodness of fit test that we executed to assess the statistical significance of subjects' level of agreement. With regard to the elicitation workshops, the chi-square test results show that the subjects' moderate agreement with the fact that CMC elicitations encourage to more openly discuss conflicting issues with same and other group members (item 2 and 3, respectively) is significant at the 5% level ($\chi^2=11.48$, $p=.009$, and $\chi^2=9.12$, $p=.028$, respectively). With respect to the negotiation workshops, the chi-square test results show that subjects' moderate agreement with having increased opportunity to participate in the discussion and being encouraged to more openly discuss conflicting issues with same group members during CMC negotiations (item 1 and 3, respectively) is significant at the 5% level ($\chi^2=10.68$, $p=.014$, and $\chi^2=8$, $p=.018$, respectively).

In general, the results of the goodness of fit tests show the subjects tending to somewhat agree that, compared to F2F requirements workshops, in CMC elicitations and negotiations they had increased opportunity to participate and more openly discuss about conflicting issues with the other participants.

These statistics, however, compared F2F elicitation to CMC elicitation, and F2F negotiation to CMC negotiation through subjects' responses, regardless of the fact that they participated in either requirements workshop playing different roles. Hence, we performed a t-test to verify whether being client or developer influenced subjects' perception of comfort with communication mode in both paired comparisons. As a nonparametric alternative to t-test on independent samples, we applied the Mann-Whitney U test [5], but we failed to find any significant difference.

Table 3. Results from the Wilcoxon signed-rank test for the matched-pair comparisons

Matched-pair comparison A vs. B	Positive ranks A > B	Negative ranks A < B	Ties A = B	Wilcoxon test
I. F2F elicitation vs. F2F negotiation	0	6	4	**Z=2.27**
II. F2F elicitation vs. CMC elicitation	5	2	3	Z=1.56
III. F2F negotiation vs. CMC negotiation	8	0	2	**Z=2.54**
IV. CMC elicitation vs. CMC negotiation	4	3	3	Z=.88

Results significant at the 5% level are shown in bold

Table 4. Evaluation of comfort with communication mode and results from the chi-square goodness of fit test

"In comparison to F2F workshops, CMC workshops…"

	1. "offered increased opportunity to participate in the discussion"				***2. "encouraged to more openly discuss conflicting issues with same group members"***				***3. "encouraged to more openly discuss conflicting issues with other group members"***			
	Elicitation		**Negotiation**		**Elicitation**		**Negotiation**		**Elicitation**		**Negotiation**	
	N	%	N	%	N	%	N	%	N	%	N	%
Strongly agree	5	29.4	1	5.3	2	11.8	2	10.5	5	29.4	-	-
Somewhat agree	6	35.3	**10**	**52.6**	**10**	**58.8**	8	42.1	**9**	**52.9**	**11**	**57.9**
Somewhat disagree	4	23.5	6	31.6	4	23.5	6	31.5	2	11.8	7	36.8
Strongly disagree	3	11.8	2	10.5	1	5.9	3	15.8	1	5.9	1	5.3
χ^2	2.06		**10.68**		**11.48**		4.79		**9.2**		**8**	

Significant results at the 5% level are shown in bold

4. Discussion

In this study we compared the use of F2F and synchronous, text-based communication (CMC) for supporting ad hoc groups of stakeholders involved in distributed requirements workshops. In the comparison we evaluated the levels of *comfort with communication mode* and *satisfaction with performance* perceived by stakeholders. Because the role factor was not found to significantly affect our findings, it is ignored in this discussion.

With regard to the comfort with communication mode, the absolute comparison between F2F and CMC is not under investigation: The prominent theories of media richness have already acknowledged the general individual preference for rich interaction, regardless of any context [7],[26]. However, our findings in the context of distributed Requirements Engineering show that the stakeholders significantly perceived to have increased opportunity to participate and more openly discuss conflicting issues with other participants during CMC elicitations and negotiations, compared to F2F workshops (see Table 4). These results confirm the predictions of the socio-psychological theories, which argue that the depersonalization effect induced by the use of less-rich and less-social media limits domination, group/social pressure, and the other dysfunctional aspects intrinsic to F2F group communication [28].

GSS research has shown that groups interacting on text-based channels have often outperformed collocated groups in task of idea generation because of the possibility to input ideas in parallel. Conversely, collocated groups have usually outperformed distributed groups in executing tasks that involve problem solving, decision making, and conflict resolution [21]. Neither the use of rich media, like video or F2F communication, has been shown to positively affect the performance quality of the work when it involves negotiation [11],[24]. Thus, consistently with these findings, we expect that synchronous, text-based elicitation represents a better task/technology fit than synchronous, text-based negotiation. The box plot in Figure 2 shows a large and statistically significant difference between subjects' satisfaction with performance during F2F and CMC negotiations, perceived as the best and worst fit, respectively. In contrast, the difference between F2F and CMC elicitation is not statistically significant. These results, on the one hand, confirm that in terms of satisfaction with performance CMC elicitation is a better task/technology fit than CMC negotiation, and, on the other hand, suggest that the general preference for F2F requirements workshops is due to the strong preference for the F2F negotiation fit over the CMC counterpart (see Table 3).

Nevertheless, further analyses are needed to provide a more thorough answer, since these results only address stakeholders' perception of the workshop process performance. Further evidence, not related to subjects' perceived satisfaction, can be obtained analyzing the differences in the workshop outcomes,

that is, the specification documents RS 1.0 and RS 2.0, produced by stakeholders as a result of the requirements elicitations and negotiations. The *requirements workshop* and *communication mode* factors created two variants in the iterative process used to produce the requirements specification document (see Figure 1). In our study three of the six projects were completed following the first process variant, which includes CMC elicitations and F2F negotiation workshops. Conversely, the remaining three projects were completed following the second process variant, which includes instead, F2F elicitation and CMC negotiation workshops. Hence, to confirm that CMC elicitation represents a better task/technology fit, the results from the analyses of the specification documents are expected to show that (1) no significant difference is found with respect to the quality of the RS 1.0 (e.g., completeness, correctness), whatever the process variant utilized; (2) the RS 2.0 documents created following the first process variant (CMC elicitation and F2F negotiations) are better than those created following the second process variant (F2F elicitation and CMC negotiation).

5. Threats to Validity

One of the key issues in experimentation is evaluating the validity of results [29]. In this section the validity of the findings is discussed with respect to the threats that are relevant for our study.

Threats to internal validity influence the conclusions about a possible causal relationship between the treatment and the outcome of a study. The following rival explanations for the findings have been identified.

Because we evaluated the interaction between stakeholders who defined the requirements for six different applications, it cannot be excluded that the differences in the application domain and complexity have influenced our study as confounding factors.

Since our experiment was performed during a considerable part of the whole course, boredom and tiredness effects cannot be disregarded and may partially explain the lower return-rate of the post-negotiation (second) questionnaire. However, since the subjects were graded on the overall outcome of the requirements definition process (i.e., the RS 2.0), they were motivated to keep a deep commitment to the tasks.

External validity describes the study representativeness and the ability to generalize the results outside the scope of the study. We identified the following threats to external validity.

For any academic laboratory experiment the ability to generalize results to industry practice is restricted by the employment of students as study participants. While the students may not be representative of the entire population of software professionals, it has been shown that the differences between students and real developers may not be as large as assumed by previous research [15]. Another issue with the representativeness of subjects is related to their familiarity with the use of synchronous, text-based communication. Computer science students are very accustomed with text-based interaction. Nevertheless, synchronous, text-based communication tools, such as chat and IM, are increasingly being adopted in the workplace, not only in the field of software development, to complement email [14].

Another threat to generalizeability is the simulation of the geographical dispersion. The subjects in the study were not actually dispersed. Instead, the members of each team were distributed in the three laboratory used during the CMC requirements workshop sessions. This threat was mitigated to some extent by the strict control asserted over the students in order to prevent them from interacting verbally throughout the workshops. Nevertheless, we could prevent rich interaction between the subjects only during the CMC sessions, while it cannot be excluded that the students had follow-up F2F discussions after.

Construct validity concerns the degree of accuracy to which the variables defined in the study measure the constructs of interests. We identified a couple of threats to construct validity.

The constructs of *satisfaction with performance* and *comfort with communication mode* have been adapted from a similar study on media effects [21]. The several questions used to measures these construct were defined taking into account (1) the communication issues commonly experienced and already acknowledged by previous research in requirements engineering, and (2) the issues informally reported by the students. While one could argue about the arbitrariness in the definition of the scales used to operationalize each construct, in our study this issue has been overcome by executing principal component analysis and scale reliability analysis.

Finally, our measures of the constructs are taken from self-reported data. However, subjects' preference for the communication mode is not always aligned with actual performance gaining, as shown by GSS-related research. In our study this drawback is mitigated by having the subjects express their media preference not for hypothetical situations, but upon the accomplishment of realistic experimental tasks (i.e., the requirements workshops).

6. Conclusions

In this paper, we have presented an experiment conducted at the University of Victoria on the effects of synchronous, text-based communication in distributed requirements workshops. In particular, we analyzed the differences between F2F and text-based communication in terms of satisfaction with performance and comfort with communication mode, as perceived by stakeholders during both elicitation and negotiation workshops.

Differently from many other experiments on media effects, this study did not use generic, puzzles-like tasks that involve either idea generation or problem solving. Instead, the experimental tasks were elicitations and negotiations of software requirements for non-toy-sized, realistic systems. In addition, the participants needed to recall specialized knowledge (e.g., the RFP during the elicitation workshops, the RS 1.0 during the negotiations) and techniques learned through the course (e.g., meeting facilitation), to effectively accomplish the tasks. This resulted in a higher cognitive load for the study participants and an increased, more realistic effort required for accomplishing the experimental tasks [21].

The findings from the first analyses of the experimental data have confirmed, on the one hand, that in terms of satisfaction with performance CMC elicitation is a better task/technology fit than CMC negotiation, and, on the other hand, that the general preference for F2F over CMC is due to the strong preference for the F2F negotiation fit over the CMC counterpart. These findings resulted from the analysis performed on the post-hoc satisfaction questionnaires administered to the subjects after the requirements workshops and hence, they specifically address how stakeholders perceived the workshop process performance. Nevertheless, in order to accurately assess the effectiveness of using a synchronous, text-based communication channel for conducting requirements workshops, we need to perform further analyses on data other than those self-reported on satisfaction questionnaires. Indeed, the large body of knowledge about media richness has proved that asking directly about both media preferences and media effectiveness is deeply affected from the perceived richness and social presence of the media themselves, regardless of the type of task [26], [7], [30]. Thus, to confirm our findings, we are analyzing the differences in the workshop outcomes, that is, the specification documents RS 1.0 and RS 2.0, produced by stakeholders as a result of the requirements elicitation and negotiation workshops, respectively.

Acknowledgments

We would like to thank Luis Izquierdo and Irwin Kwan for their support, and all the students who took part in the experiment.

References

[1] Al-Rawas, A., and Easterbrook, S. "Communication Problems In Requirements Engineering: A Field Study." *Proc. 1st Conf. on Professional Awareness in Software Engineering*, Royal Society, London, 1-2 Feb., 1996.

[2] Birnholtz, J.P., Finholt, T.A., Horn, D.B., and Bae, S.J. "Grounding Needs: Achieving Common Ground Via Lightweight Chat In Large, Distributed, Ad-Hoc Groups." *Proc. Int'l Conf. on Human Factors in Computing Systems (CHI '05)*, Portland, USA, April 2-7, 2005, pp. 21-30.

[3] Calefato, F., and Lanubile, F. "Using The eConference Tool for Synchronous Distributed Requirements Workshops", *Proc. 1st Int'l Workshop on Distributed Software Development (DiSD '05)*, Paris, France, Austrian Computer Society, 29 August, 2005, pp. 97-108

[4] Carlson, J.R., and Zmud, R.W. "Channel Expansion Theory and The Experiential Nature of Media Richness Perceptions." *Academic Management Journal*, Vol. 42, No. 2, 1999, pp. 153–170.

[5] Conover, W.J. *Practical Nonparametric Statistics.* J. Wiley & Sons, New York, NJ, 1980.

[6] Cronbach, L.J. "Coefficient Alpha and The Internal Structure of Tests." *Psychometrika,* Vol. 16, 1951, pp. 297-334.

[7] Daft, R.L., and Lengel, R.H. "Organizational Information Requirements, Media Richness and Structural Design." Management Science, Vol. 32, No. 5, 1986, pp. 554-571.

[8] Damian, D., and Zowghi, D. "Requirements Engineering Challenges in Multi-Site Software Development Organizations." *Requirements Engineering Journal*, Vol. 8, 2003, pp. 149-160.

[9] Davis, P.K, and Wainfan, L. *Challenges In Virtual Collaboration: Videoconferencing, Audioconferencing, and Computer-mediated communication*. Rand Corporation, 2004.

[10] Dennis, A.R., and Valacich. J.S. "Testing Media Richness Theory in the New Media: The Effects of Cues, Feedback, and Task Equivocality." *Information Systems Research*, Vol. 9, No. 3, 1998, pp. 156-274.

[11] Finn, K.E., Sellen, A.J., and Wilbur, S.B. *Video-Mediated Communication*. Lawrence Erlbaum Assoc. Inc., Hillsdale, NJ, 1997.

[12] Fjermestad, J., and Hiltz, R. Case and Field Studies of Group Support Systems: An Empirical Assessment. *Proc. Int'l Hawaii Conference on System Science (HICSS-33)*, Hawaii, Jan. Vol. 1, pp. 4-7.

[13] Hatcher, L., and Stepanski, E.J. *Step-By-Step Approach to Using the Sas System for Univariate and Multivariate Statistics*, Paperback, 1994.

[14] Herbsleb, J.D., Atkins, D.L., Boyer, D.G., Handel, M., and Finholt, T.A. "Introducing Instant Messaging and Chat into the Workplace." *Proc. Int'l Conference on Computer-Human Interaction (CHI '02)*, Minneapolis, MN, USA, 2002.

[15] Höst, M., Regnell, B., and Wohlin, C. "Using Students as Subjects - A Comparative Study of Students and Professionals in Lead-Time Impact Assessment." *Empirical Software Engineering*, Vol. 5, No. 3, 2000, pp. 201-214.

[16] Isaacs, E.A., and Tang, J.C. "What Video Can and Cannot Do for Collaboration: A Case Study." *Multimedia Systems*, Vol. 2, 1994, pp. 63-73.

[17] Lloyd., W.J., Rosson, M.B., and Arthur, J.D. "Effectiveness of Elicitation Techniques in Distributed Requirements Engineering." *Proc. IEEE Int'l Conf. on Requirements Engineering (RE '02)*, Essen, Germany, 9-13 September 2002, pp. 311- 318.

[18] Mark, G. , Grudin, J., and Poltrock, S.E. "Meeting at the desktop: an empirical study of virtually collocated teams. *Proc. 6th European Conf. on Computer supported cooperative work (ECSCW '99)*, Copenhagen, Denmark, 1999.

[19] Macaulay, L.A. *Requirements Engineering*. Springer-Verlag Telos, 1996.

[20] Montgomery, D.C. *Design and Analysis of Experiments*. J. Wiley & Sons, New York, 1996.

[21] Murthy, U.S., and Kerr, D.S. "Task/Technology Fit and The Effectiveness of Group Support Systems: Evidence in The Context of Tasks Requiring Domain Specific Knowledge." *Proc. 33rd Hawaii Int'l Conf. on System Sciences (HICSS-33)*, 2000, Vol. 2, pp. 1-10.

[22] Nunnally, J. *Psychometric Theory*. McGraw-Hill, New York, NJ, 1978.

[23] Olson, J.S., Olson, G.M., and Meader, D.K. "What Mix of Video and Audio is Useful for Remote Real-Time Work?" *Proc. Int'l Conf. on Human Factors in Computing Systems (CHI '95)*. Denver, USA, 1995, pp. 362-368.

[24] Olson, J.S., Olson, G.M., and Meader, D. "Face-To-Face Group Work Compared to Remote Group Work with and without Video." In Finn, K., Sellen, A., and Wilbur, S. (eds.), *Video Mediated Communication*, Hillsdale, NJ: Lawrence Erlbaum Associates, 1997.

[25] Poltrock S.E., and Grudin J. "Videoconferencing: Recent Experiments and Reassessment." *Proc. 38th Hawaii Int'l Conf. on System Sciences (HICSS-38)*, 2005, pp. 104a-104a.

[26] Short, J., Williams, E., and Christie, B. *The Social Psychology of Telecommunications*. London: John Wiley and Sons, 1976.

[27] Tang, J.C., and Isaacs, E.A. "Why Do Users Like Video?" *Computer Supported Cooperative Work*, Vol. 1, 1993, pp. 163-196.

[28] Walther, J.B. "Computer-mediated Communication: Impersonal, Interpersonal, and Hyperpersonal Interaction." *Communication Research*, Vol. 23, No.1, 1996, pp. 3-43.

[29] Wohlin, C., Runesson, P., Höst, M., Ohlsson, M.C., Regnell, B., and Wesslén, *A. Experimentation in Software Engineering, An Introduction*. Kluwer Academic Publishers, 2000.

[30] Zigurs, I. and Buckland, B.K. "A Theory of Task/Technology Fit and Group Support Systems Effectiveness." *MIS Quarterly*, Vol. 22, No. 3, Sep. 1998, pp. 313-334.