An approach to assess the quality of collaboration in technology-mediated design situations

Jean-Marie Burkhardt
Paris Descartes University,
Ergonomics–Behavior-Interaction Lab.
45 rue des Saints-Pères
75270 Paris Cedex 06, France
jean-marie.burkhardt@univ-paris5.fr

Françoise Détienne[1]
LTCI- UMR 5141
CNRS - Telecom ParisTech
46 rue Barrault
75634 Paris Cedex 13, France
[1] INRIA-Paris France
francoise.detienne@telecom-paristech.fr

Anne-Marie Hébert
LTCI- UMR 5141
CNRS - Telecom ParisTech
46 rue Barrault
75634 Paris Cedex 13, France
hebert@telecom-paristech.fr

Laurence Perron
France Telecom R&D / Orange Labs
Lannion, France
laurence.perron@orange-ftgroup.com

Stéphane safin
Lucid Group
Université de Liège –Belgique
stephane.safin@ulg.ac.be

Pierre Leclercq
Lucid Group
Université de Liège –Belgique
pierre.leclercq@ulg.ac.be

ABSTRACT

Our objective is to measure and compare the quality of collaboration in technology-mediated design activities. Our position is to consider collaboration as multidimensional. We present a method to assess quality of collaboration which is composed of seven dimensions concerning communication processes such as grounding, coordination processes, task-related processes, symmetry of individual contributions as well as motivational processes. This method is used in a study aiming to compare the quality of collaboration in architectural design. In this experimental study, design situations vary according to technology-mediation - co-presence with an augmented reality (AR) environment versus distance with AR and visio-conferencing -, and according to number of participants - pairs versus groups of four architects -. Our results show that distinctive dimensions of collaboration are affected by the technology mediation and/or the number of co-designers. We discuss these results with respect to technology affordances such as visibility and group factors.

Keywords

collaboration, design, methodology, cognitive ergonomics, Computer Supported Collaborative Learning.

ACM Classification Keywords

C4. Design Sudies. H.5.2 User Interface. ergonomics.

1. INTRODUCTION

With the growing importance of technology mediation for group work, developing methods for assessing the quality of collaboration should become as central as developing methods for assessing usability of UI in user-centred design. In spite of a growing number of methods to evaluate groupware technologies and group work, no measurement method of this facet of collaboration has been proposed as far as we know.

Our objective is to measure and compare the quality of collaboration in technology-mediated design activities. In this context, the term 'quality' can be understood in descriptive terms (identifying and discriminating the intrinsic properties of collaboration) and/or in a normative sense (identifying what makes 'good' or less good collaboration, considered *sui generis*). We consider these visions of quality as complementary. On the basis of previous work in cognitive ergonomics of design and in computer supported collaborative learning (CSCL), we consider collaboration as multidimensional and propose an evaluation method that covers these specific dimensions. Furthermore, whenever possible we provide norms and explicit qualitative references to support the comparison of measures in various technology-mediated situations.

In the first part of the paper, we provide the rationale of this multidimensional approach on the basis of theoretical arguments and of results from empirical studies. We also aim to elicit references (often implicit in the literature) and relevant standards regarding collaborative activities. A method to assess selected dimensions of activity related to collaboration quality is then presented, followed by the test of its reliability. In the second part, we present a study aiming to compare the quality of collaboration, relatively to our various dimensions, in contrasted technology-mediated design situations. We use our assessment method to compare quality of collaboration in design situations varying according to technology-mediation - co-presence with an augmented reality (AR) environment *versus* distance with AR and visio-conferencing - and according to number of participants - pairs *versus* groups of four architects. The results of this empirical study are presented and discussed.

2. COLLABORATION IN DESIGN: A MULTI-DIMENSIONAL APPROACH

Empirical studies on the process of collaboration in design teams (for a state of the art, see [1]), in various application domains (e.g., software design, architectural design), have highlighted distinctive collaborative processes most important for successful design. These processes can be taken as a referential of good collaboration with respect to design. They can be grouped along several dimensions concerning communication processes such as grounding, task-related processes (e.g. exchanges of knowledge relevant for the task at hand; argumentation processes), coordination processes, and motivational processes. Furthermore we consider, through all these dimensions, how symmetric individual contributions are in order to provide a complementary aspect of collaboration and its quality.

2.1 Communication processes

Communication processes are most important to ensure the construction of a common referential within a group of collaborators. The establishment of common ground is a collaborative process [2] in which the co-designers mutually establish what they know so that design activities can proceed. Grounding is linked to sharing of information through the representation of the environment and the artefact, the dialog, and the supposed "pre-existing" shared knowledge. This activity ensures inter-comprehension and construction of shared (or compatible) representations of the current state of the problem, solutions, plans, design rules and more general design knowledge.

Empirical studies of collaborative design (e.g. [3, 4, 5]) found that grounding, although time-consuming, was most important to ensure good design: for instance, Stempfle and Badke-Schaub [5] found that when teams bypassed grounding (referred to as "analysis"), this led them to premature evaluation of design ideas.

Key characteristics of collocated synchronous interactions are assumed to support grounding [6]. Rapid feedbacks allow for rapid corrections when there are misunderstandings or disagreements. Multiple channels (visual, oral, etc.) allow for several ways to convey complex messages and provide redundancy: e.g., gaze and gestures help to easily identify the referent of deictic terms. The shared local context allows for mutual understanding. At distance, characteristics of communication media, such as the lack of visibility or simultaneity [2], may affect grounding and awareness. Using videoconferencing tools can extend the channels by which people communicate.

2.2 Task-related processes

Task-related processes concern the evolution of the design problem and solution: (a) design activities, i.e., elaboration, enhancements of solutions and of alternative solutions; (b) evaluation activities, i.e., evaluation of solutions or alternative solutions, on the basis of criteria. These activities are supported by argumentation and negotiation mechanisms. These content-oriented mechanisms reveal how the group resolves the task at hand by sharing and co-elaborating knowledge concerning the design artefact, by confronting their various perspectives, and by converging toward negotiated solutions.

Whereas studies show evidence that these mechanisms are important for the quality of design products (e.g. [7]), empirical studies show that important drawbacks of observed design teams (e.g. [5]) may be: limitation in solution search; early choice of a solution without exploration of all alternatives; rapid solution evaluation on the basis of just a few criteria; difficulties in taking into account all criteria and their inter-dependencies (constraint management).

Technology-mediation tends to have few effects on these design processes. For example, a previous study [8] showed a similar distribution of the main categories (based on a coding scheme of interactions) of activities related to design for pairs of architects in co-presence and at distance (with videoconferencing and digital tablets). This absence of effect could be, however, specific to synchronous collaborative situations.

2.3 Group management processes

Collaboration concerns group management activities such as: (a) project management and coordination activities, e.g., allocation and planning of tasks; (b) meeting management activities, e.g., ordering, postponing of topics in the meeting. These process-oriented mechanisms ensure the management of tasks interdependencies, which is most important in a tightly coupled task such as design. These coordination mechanisms tend to become more central with technology mediation [e.g, 9].

2.4 Cooperative orientation and motivation

Although less covered in previous studies on design as well as in studies on technology-mediated collaborative design, cooperative orientation and motivation may be considered as important aspects of collaboration. Indeed, recent research on collaboration processes in design [10, 11] considers the participants' roles in communication, group management and task management. The balance between these roles is considered as a good indicator of collaboration. This aspect is similar to the notion of reciprocal interaction highlighted by Spada et al. [12] and symmetry in the interaction pointed out by Baker [13] or Dillenbourg [14] in CSCL. These authors consider that quality of collaborative learning in small groups of learners is linked to the symmetry of the interaction. We will adopt a similar posture to assess the quality of collaboration in small teams of designers. We will also consider the dimension of motivation as important in so far as it can strongly affect the actual way of collaboration.

3. CURRENT APPROACHES OF COLLABORATION IN EVALUATION METHODS OF GROUPWARE TECHNOLOGIES

Groupware evaluation methods are of two types: usability inspection methods and usability evaluation methods based on users' studies. On the inspection side, usability evaluation techniques [e.g., 15] do not rely on the participation of real users of the system. Advocates of these techniques argue that they are less costly than field methods and they can be used earlier in the development process. However their focus remain on individual-centered task models [e.g., 16], i.e. eliciting goals and actions required for users to interact together and not on the collaboration processes and their quality *per se*. Furthermore, they do not explore effective collaboration processes in context.

Regarding user studies, there is a lot of methods which rely on different data collection and analyses techniques: they can be based on computers logs, interactions between participants (coding methods or ethno-methodological methods), or interviews. In field or experimental studies, the few indicators used to assess usability regarding collaboration processes are focused on quantifying fine-grained interactions. An example, given in a recent review by Hornbaek [17], concerns the measure of “communication effort”: number of speakers’ turns; number of words spoken; number of interruptions; amount of grounding questions. Furthermore, quantitative variations of such indicators are non-univocal: any increase or decrease of them could signify either an interactive-intensive collaboration or evidence of huge difficulties in establishing or maintaining the collaboration. Several other drawbacks of these methods are usually pointed out: they are often difficult (and sometime impossible) to apply with prototypes and they are most time-consuming.

Two additional criticisms are the extent to which existing empirical-based methods cover all the dimensions of collaboration and their generality or *ad hoc* nature. Indeed, user-based methods to assess collaboration usually concentrate only on one or two dimensions among the numerous ones we wish to cover: for example, verbal and gestural communication to assess the grounding processes. Furthermore, motivational aspects as well as the balance/symmetry of individual contributions are rarely considered, although they reflect complementary aspects of collaboration assessment. Assessment methods also vary according to their generality and their explicit/formal characteristics. Ethno-methodological approaches generally remain *ad hoc* to the analyzed situations and they do not rely on any explicit methods. This is based on the adopted theoretical position considering the particular context under study as most important. Other user-based approaches rely on coding schemes making explicit categories of analysis, but they often remain *ad hoc* to the observed situation. Still, in some task application domains, some efforts have been made to construct more generic categories (see for example, [4, 5, 8]). To summarize, none of the user-based methods[1] developed in the Computer-Supported Cooperative Work (CSCW) field are both multidimensional and generic.

In the close field of CSCL, the analysis of the process of collaboration is also a central topic of research. The Spada rating scheme [12, 18, 19] is certainly the most representative of recent effort made in this field to assess collaboration and its quality. It has been developed to compare and assess collaboration in collaborative learning tasks, with respect to various learning methods or technical support. An advantage of this method, beside its low temporal cost[2], is its coverage of a wide range of collaboration dimensions. Indeed, these authors consider nine qualitatively defined dimensions that cover five broad aspects of the collaboration process: communication (sustaining mutual understanding, dialogue management), joint information processing (information pooling, reaching consensus), coordination (task division, time management, technical coordination), interpersonal relationship (reciprocal interaction) and motivation (individual task orientation). A review of the literature [19] on computer-supported collaborative learning and working provides theoretical arguments to consider these five aspects as central for successful collaboration under the conditions of video-mediated communication and complementary expertise. Furthermore, their method is generic enough to be applied to different technology-mediated situations.

However, Spada’s method shows some limits from our viewpoint. Indicators exploited by judges (or raters) in order to assess collaboration are underspecified. Indeed, the method relies essentially on the subjective evaluation of each given dimension on a 5-grade Likert-like scale, oriented by a training manual. One consequence is that such an approach hides the reference to any quantifiable events or observables from the collaborative situation, preventing any track back from the assessment values to original data. However, its multidimensional characteristic endows the method with a good basis to further develop a method to assess the quality of collaboration in technology-mediated design. In this objective, we modified the assessment procedure to make the observable indicators underlying the evaluation more explicit. This is reported in the next section presenting our method.

4. THE MULTI-DIMENSIONAL METHOD

We propose a multi-dimensional method to evaluate the quality of collaboration in technology-mediated design. Our method is initially (and thus partly) based on Spada’s method [12]. It has been modified so as to take into account characteristics of collaborative design and to improve the assessment procedure.

4.1 Assessment procedure, dimensions and indicators of collaboration

In an initial version, our method kept the principle of subjective scale rating by Spada , i.e., the judges were requested to give a score on a 5-grade Likert scale for each of the dimensions. However to generate explicit traces of the rating processes, we modified the scoring method by requesting the judges to give additional explicit answers (Yes, Yes/No, No) to questions related to the specific indicators of each dimension (Table 1). For each indicator, we balanced questions with positive valence and questions with negative valence. These questions distinguish between what we consider as “good” collaboration (question with position valence) and collaboration with a lowest quality (questions with negative valence) with respect to successful collaborative design. As an illustration, let us consider two examples. For the indicator “mutual understanding of the state of design problem/solutions” of *Dimension 2 (Sustaining mutual understanding)* the judge is requested to answer two questions (by Yes, Yes/No, or No): the question with positive valence is “Do the designers ask questions, give clarifications or complementary information, using verbal or behavioural backchannels, on the state of the design artefact?”. The question with negative valence is “Are there misunderstandings on the state of the design artefact during relatively long periods of time?”. As another example, the indicator “common decision taking” of *Dimension 4 (Argumentation and reaching consensus)* is splitted up into two questions: the question with positive valence is “Are the individual contributions equal concerning the design choices?”; the question with negative valence is “Is there one contributor who imposes the design choices?”.

Consequently, assessment of the quality of collaboration was based on the rating of the seven dimensions plus 46 questions distributed along these dimensions. There were 3 positive + 3 negative questions for dimensions 1, 2, 3, 4, 7 and 4 positive + 4

---

[1] Modeling techniques such as CUA (Collaboration Usability Analysis) [15] could be considered as multidimensional and generic. However this method does not enter into the category of user-based methods.

[2] The judges can directly apply the rating method on the basis of video recordings. Thus they do not rely on transcriptions which are time-consuming.

negative questions for dimensions 5, 6. For dimension 7, both the questions and the scale rating are applied to each participant of the assessed collaborative situation.

Table 1. Dimensions and indicators of our method

Dimensions	Definition	Indicators
1. Fluidity of collaboration	It assesses the management of verbal communication (verbal turns), of actions (tool use) and of attention orientation.	– Fluidity of verbal turns – Fluidity of tools use (stylet, menu) – Coherency of attention orientation
2. Sustaining mutual understanding	It assesses the grounding processes concerning the design artefact (problem, solutions), the designers' actions and the state of the AR disposal (e.g., activated functions).	– Mutual understanding of the state of design problem/solutions – Mutual understanding of the actions in progress and next actions – Mutual understanding of the state of the system (active functions, open documents)
3. Information exchanges for problem solving	It assesses design ideas pooling, refinement of design ideas and coherency of ideas.	– Generation of design ideas (problem, solutions, past cases, constraints) – Refinement of design ideas – Coherency and follow up of ideas
4. Argumentation and reaching consensus	It assesses whether or not there is argumentation and decision taken on common consensus.	– Criticisms and argumentation – Checking solutions adequacy with design constraints – Common decision taking
5. Task and time management	It assesses the planning (e.g. task allocation) and time management.	– Work planning – Task division – Distribution and management of tasks interdependencies – Time management
6. Cooperative orientation	It assesses the balance of contribution of the actors in design, planning, and in verbal and graphical actions.	– Symmetry of verbal contributions – Symmetry of use of graphical tools – Symmetry in task management – Symmetry in design choices
7. Individual task orientation	It assesses, for each contributor, its motivation (marks of interest in the collaboration), implication (actions) and involvement (attention orientation).	– Showing up motivation and encouraging others motivation – Constancy of effort put in the task – Attention orientation in relation with the design task

We also modified the definition of dimensions to take into account the nature of the design task and the technology mediation characteristics[3]. For the first aspect, we made explicit design task processes for *Dimension 3 (Information exchanges for problem solving)* and *Dimension 4 (Argumentation and reaching consensus)*: in particular those related to design constraints, design problem, design solutions, design decisions. For the second aspect, we distinguished for *Dimension 2 (Sustaining mutual understanding)* between processes related to the state of the design, the state of the system, and actions in progress.

4.2 Testing the reliability of the method

We tested the reliability as well as the usability of an initial version of the method on the basis of inter-raters correlations, interviews and analyses of judges' activity during their application of the method (see [20] for the evaluation approach). Four judges participated in this first test (one of them is the third author of this paper). After an initial training on the method involving its application on one excerpt of 15 minutes, followed by a debriefing, the raters were provided with three excerpts of 10 minutes to be assessed. The test of this early version showed a weak inter-raters agreement of subjective rating, i.e., on 5-grade notes on dimensions. Conversely, we observed a high inter-raters agreement in the way they responded to the specific questions (all kappas > 0.60). Additionally, 17% of the questions were either misunderstood or judged as not applicable to the extracts.

On the basis of these results, the difficult questions were either reformulated or withdrawn and we decided to adopt an explicit scoring algorithm based on the number of positive and negative answers to questions, instead of subjectively rating each dimension. Thus, the score for each dimension is calculated on the basis of the answers given by the judge to the questions. The reliability of this second version was tested by looking again at inter-raters agreement. Three judges (the first three co-authors of this paper) assessed independently the quality of collaboration in an excerpt of collaborative design task. After a initial training on the method involving its application on two

[3] We also added items to make explicit the collaborative modalities (verbal, gestural, graphical, textual) predominantly associated to each considered dimensions of collaboration. We will not discuss this aspect of the evaluation in this paper.

excerpts of 15 minutes, each followed by a short debriefing to elicit and share a set of common rules with the method, the three raters were provided the same (and previously not viewed) excerpt to be assessed. The analysis of all the responses to questions for all dimensions shows an excellent inter-rater agreement (96.15% of agreement between the three raters on all questions; Kappa=.92). Furthermore, specific analyses of responses show an excellent agreement for both Yes (Kappa=.94) et No answers (Kappa=.92). Only Yes/No questions exhibited a lack of agreement (Kappa=0.00) due to their very low frequency (2/78). This can also be explained by a constructed assessment rule which was to avoid this answer mode as far as possible and to favour a clear Yes or No decision. We used this version of method in the empirical study presented in the following section of this paper.

5. EMPIRICAL STUDY OF QUALITY OF COLLABORATION IN TECHNOLOGY-MEDIATED DESIGN SITUATIONS

We conducted an empirical study to compare the quality of collaboration with our method across three distinctive technology-mediated design situations. The design domain is architecture. Briefly, groups of two or four architects were asked to collaborate to solve a design task, either in co-presence around a virtual desktop or at distance with a virtual desktop on each site and a videoconferencing set up. The work sessions were entirely videotaped and solutions recorded. They represent about 18h of video records.

5.1 Participants

Sixteen last year students in architecture or in building engineering and architecture (from 22 to 26 years old) participated in this study. They had a similar experience in design tasks, to avoid biases due to the diversity of professional practices, particularly important in architecture. They were distributed arbitrarily into three conditions: pairs in co-presence, pairs at distance and 4-designers groups at distance. They were paid for their participation.

5.2 The three technology-mediated experimental conditions

The study is based on an integrated aided design tool, the EsQUIsE software, based on a Virtual Desktop [21, 22]. This environment has been developed to assist architects in preliminary design while keeping the natural and simple characteristics of the pen/paper drawing process. It is composed of a mixed software and hardware solution which offers (i) the natural aspect of digital freehand sketching, (ii) the ability of drawing interpretation and generation of 3D models based on 2D sketches and, (iii) the direct model manipulation and evaluation of performances (presently in building engineering). The system consists in a classical “A0” desk with a suspended ceiling equipped with a double projection system offering a large working surface (approximately 150x60 cm). The electronic stylus allows drawings of virtual sketches on this surface. The designers can manipulate their drawings and are provided with automatically generated models without having to use any usual modeller in the AR environment. To ensure the sharing of data in real time at distance, the software Sketsha was used in the distant conditions. This software is very similar to EsQUIsE except for the sketches interpretation and the generation of 3D models, which are not supported.

We contrasted three technology-mediated situations with two observations for each situation:

- Two pairs were in co-presence (cf. Figure 1): architects of a pair were sat side by side on a virtual desktop.
- Two pairs were at distance (cf. Figure 2): each architect of a pair had at his/her disposal a virtual desktop and a videoconferencing set up.
- Two groups of 4 architects at distance were spatially distributed as follows (cf. Figure 3): each group was composed of two collocated pairs working on a virtual desktop with a videoconferencing set up.

The videoconferencing system (IChat on 17’ monitors with integrated webcam) allowed distant designers to collaborate through multiple channels and modalities (video, audio). The two virtual desktops in the distant situations were connected in a completely synchronous way.

4

Figure 1. A pair of architects collaborating in co-presence

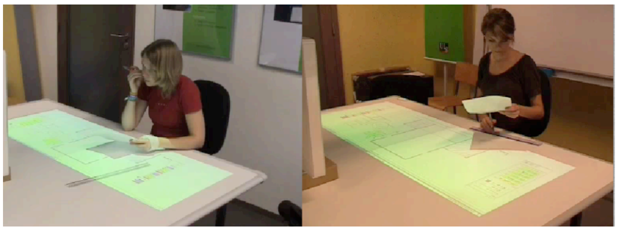

Figure 2. A pair of architects collaborating at distance

Figure 3. A Group of four architects collaborating at distance

5.3 Design problems

One version of two similar architectural design problems (rural school and urban school) was given to each pair or group. This problem is a pedagogical exercise representative of design problems commonly faced by architects in their practices [23]. Furthermore, it is assumed that the two versions of the problem, with similar constraints and proportions, do not influence the nature of the activity and the collaboration of designers.

5.4 Collected data

The experimental sessions were recorded between May 2006 and April 2008 in Liège. In the co-presence condition, we used two video-recorders to capture both a detailed view of the workspace and a larger view of interactions between participants. In the distant conditions, pairs or groups of 4 participants, we used two video-recorders per site with the same two views adjusted. The video recordings were synchronized.

5.5 Applying the method for evaluating the quality of collaboration

Two excerpts for each session were selected: one at the end of the first part and one at the beginning of the fourth part of the session. All excerpts concerned generation of solution and collective phases. So we had a total of 12 excerpts of 15 min.

One judge (third author of this paper) evaluated the quality of collaboration in each observed session, on the basis of two excerpts for each session. This evaluation was made on the basis of the second version of our method (Table 1). The judge began by assessing all excerpts from the collocated situations, then the excerpts from the pairs of architects at distance and the excerpts of the groups of architects at distance. The judge took 7 hour 30 minutes to apply the assessment method on the 3 hours of excerpts. We also verified that all design solutions elaborated by participants satisfied the constraints of the design problem.

5.6 Results

5.6.1 Sensitivity of dimensions to experimental conditions

Scores on dimensions exhibit different sensitivity to the various cooperative situations involved in this research (Table 2). Whereas we found a minimal difference of 0.5 point (10%) between highest and lowest scores for *Dimensions 1, 3 and 5* (*Fluidity of collaboration, Information exchange for problem solving, Task and Time management*), we observed differences up to 2 points (40%) between some configurations for *Dimension 6* (*Cooperative orientation)*. The *Dimensions 2, 4, and 7* show differences that fall within this range: *Sustaining mutual understanding* (1, i.e. 20%), *Argumentation and reaching consensus* (1, i.e. 20%), *Individual task orientation* (0.8, i.e. 16%).

Comparing respective scores in the three conditions for each dimension (Table 2) shows the following patterns of results. For Dimension 4 (*Argumentation and reaching consensus*), the score decreases from pairs in co-presence to groups of four designers at distance. For three dimensions (2, 6, 7), the mixed presence-distance condition with 4 designers exhibited clearly lower scores than pairs at distance as well as collocated pairs: *Sustaining mutual understanding, Cooperative orientation*, *Individual task orientation*. Here, the number of participants – and the intertwining between distant and collocated collaboration – could have a negative effect on these dimensions. Such a pattern suggests effects like social inhibition and social laziness as classically reported by social psychology in association to the increasing number of people within a collective. Both are affecting participation in terms of effort to contribute to the common task at hand and probability of engagement of participants. The poorest score on *sustaining mutual understanding* is also consistent with empirical results showing impact of technology-mediation on grounding (e.g., [6]).

Finally, we observed a highest score in *Dimension 1 (Fluidity of collaboration)* for pairs at distance over both collocated pairs and groups of 4 designers. This pattern could reflect a strong management of interactions and use of tools. We will develop this point when discussing the fluidity of verbal turns patterns in the next section.

Table 2. Global scores for the three design situations

Dimensions	Pairs in collocation	Pairs at distance	Groups of 4 at distance	*All conditions mean (sd)*
1. Fluidity of collaboration	4,0	4,5	4,0	*4,2 (0,4)*
2. Sustaining mutual understanding	4,0	4,5	3,5	*4,0 (0,6)*
3. Information exchanges for problem solving	5,0	4,5	4,5	*4,7(0,5)*
4. Argumentation and reaching consensus	5	4,5	4	*4,5 (0,8)*
5. Task and time management	3,5	3	3	*3,2(0,8)*
6. Cooperative orientation	5	5	3	*4,3 (1,2)*
7. Indivividual task orientation	4,5	4,5	3,7	*4,1 (0,6)*
All dimensions -mean (sd)	*4,4 (0,6)*	*4,4(0,7)*	*3,7 (0,8)*	*4,1(0,5)*

Obviously, no statistical inference can be made based on these results since our objective is mostly an exploratory test of the sensibility of the method carried out by applying it on clearly contrasted technology-mediated cooperative situations. However, as reported below, these patterns are strongly consistent with results in the literature.

5.6.2 Indicators reflecting differences in the quality of collaboration at distance compared to co-presence

Based on the questions for each indicator, we found that specific indicators reflected a poorest quality of collaboration at distance compared to co-presence. These indicators are:

- *Coherency of attention orientation (Dimension 1)*
- *Mutual understanding of the actions in progress and next actions (Dimension 2)*
- *Attention orientation in relation with the design task (Dimension 7)*

Firstly, problems pointed out in attention orientation at a group level *(Dimension1),* are also pointed out at the individual level *(Dimension7)*. Secondly, this lack of congruency in attention orientation can be related to the problems observed in the construction of mutual understanding of actions. These results reflect difficulties encountered by designers to construct a shared context and to be aligned on the task at distance. We can interpret these difficulties as related to the fractured space[5] of interaction, in particular for the visibility at distance. Indeed visibility is fractured in two spaces: the visibility of the state of the design artefact on the virtual desktop and the visibility of the gestures (deictics…) and gazes on the videoconferencing display.

[5] This refers to the notion of fractured ecology introduced by Luff et al. [24].

In contrast, we found that a specific indicator reflected a best quality of collaboration at distance compared to co-presence: *Fluidity of verbal turns (Dimension 1)*. In fact, there were less verbal overlaps at distance as the management of verbal turns was made verbally, e.g., addressing distant partners by their name. By contrast, in the collocated situation the management of verbal turns was based on visual cues like gazes and verbal overlaps were more frequent.

Finally, no difference was found for task related processes such as *Coherency and follow up of ideas (Dimension 3)* and *criticisms and argumentation (Dimension 4)*. This is coherent with results in our previous study, using a coding scheme [8], showing no effect of technology mediation on the design processes.

5.6.3 Indicators reflecting differences in the quality of collaboration in groups of 4 compared to pairs

Analysing answers for each indicator, we found that specific indicators reflected a poorest quality of collaboration in the group of 4 architects compared to the pairs. These indicators are *Common decision taking (Dimension 4), Constancy of effort put in the task (Dimension 7)* and all indicators related to *Cooperative orientation* (Dimension 6): *Symmetry of verbal contributions, symmetry of use of graphical tools, symmetry in task management* and *symmetry in design choices*.

Thus the groups of 4 have more difficulties than pairs to reach common decisions which is clearly related to the number factor. The results on all indicators of *Dimension 6* reveal an asymmetry of roles when they were four participants as compared to two. This asymmetry is observed at the general level of verbal and graphical contributions as well as at the design task level (tasks management and design choices). We can wonder to which extent this asymmetry is reinforced by the technology-mediation in the groups of 4. Finally, the indicator about constancy of effort reflects unequal motivation between participants in groups of 4 compared to pairs.

Interestingly, we found that the groups of 4 at distance were rated higher than the two other situations on the following indicators: *distribution and management of task interdependencies (Dimension 5)* and *time management (Dimension 5)*. This suggests that the need for coordination, growing with the number of participants, is particularly well managed by these groups.

6. CONCLUSIONS AND PERSPECTIVES

The results of our empirical study show that our method reveals interesting differences in the quality of collaboration that can be related to technology mediation or group factors. On one hand, our method produces results which are consistent with those found in previous studies using either ethno-methodological approaches (e.g., [24]) or coding schemes (e.g., [8]): in particular for attention orientation, reflecting lack of alignment on local objectives, and mutual understanding of actions. On the other hand, our method goes further by showing asymmetry of roles and decrease of motivation for the groups of four architects at distance. In particular, indicators of *Dimension 6* (*cooperative orientation*), which cross the other dimensions, seem to be particularly consistent one to each other. However, the lack of a control situation with groups of four designers in co-presence prevents us to conclude that there is a combined effect of group number and technology mediation on the symmetry of roles.

Nevertheless, the actual patterns of results should be considered cautiously due to the limitedness of the sample of situations on which the current data are extracted. Indeed, a recurring problem is the cost of multiplying the experiments with designers, both at the level of required time and the ability to access to a sufficient number of participants. Furthermore we were not able to assess the design solutions and thus discuss the relation between collaboration processes and the performance in terms of design product. This question is not trivial as design product assessment is based on a multiplicity of criteria.

Our perspectives are as follows. To go further on the analysis of collaboration quality, we plan other experiments to vary group composition and technology mediation characteristics. The question of the relationship between quality of collaboration and quality of design as well as group efficacy will be explored. To go further on the development of our method, we plan to compare in a systematic way between the results obtained with this rating method and those obtained with more time-consuming coding methods. We will also explore to which extent this method can be used by judges from the design domain, the architectural design domain in our case. Finally, we will explore to which extent the method could be used as a reflective support for the group itself to improve its collaboration process.

7. ACKNOWLEDGMENTS

This research has been partly funded by OrangeLabs. We thank Michael Baker for his input and feedbacks on our method.

8. REFERENCES

[1] Détienne, F. 2006. Collaborative design : managing task interdependencies and multiple perspectives. Interacting With Computers. 18(1), 1-20.

[2] Clark, H. H., and Brennan, S. E. 1991. Grounding in communication. In Perspectives on socially shared cognition, L. Resnick, J.-M. Levine and S. D. Teasley , Eds., Washington, DC: American Psychological Association Press. 127-149.

[3] D'Astous, P., Détienne, F., Visser, W., and Robillard, P.N. 2004. Changing our view on design evaluation meetings methodology: a study of software technical review meetings. Design Studies, 25, 625-655.

[4] Olson, G.M., Olson, J.S., Carter, M. R., and Storrosten, M. 1992. Small Group Design Meetings: An Analysis of Collaboration. Human-Computer Interaction, 7, 347-374.

[5] Stempfle, J., and Badke-Schaub, P. 2002. Thinking in design teams - an analysis of team communication. Design Studies, 23, 473-496.

[6] Olson, G.M., and Olson, J.S. 2000. Distance matters. Human-Computer Interaction, 15, 139-178.

[7] Pahl, G., Badke-Schaub, P., and Frankenberger, E. 1999. Resume of 12 years interdisciplinary empirical studies of engineering design in Germany. Design Studies, 20(5), 481.

[8] Burkhardt, J.-M., Détienne, F., Moutsingua-Mpaga, L., Perron, L., Leclercq, P., and Safin, S. 2008. Multimodal collaborative activity among architectural designers using an augmented desktop at distance or in collocation. In Proceedings of the European conference of Cognitive Ergonomics, ECCE 2008 (Madeira, Portugal, Sept 16-18, 2008).

[9] Herbsleb, J.D., and Mockus, A. 2003. An empirical study of speed and communication in globally-distributed

software development. IEEE Transactions on Software Engineering, 29(6), 481-494.

[10] Baker, M., Détienne, F., Lund, K., and Séjourné, A. 2003. Articulation entre élaboration de solutions et argumentation polyphonique. In Proceedings of Deuxièmes Journées d'Etude en Psychologie ergonomique – EPIQUE (Boulogne-Billancourt, France, 2-3 october, 2003) . 235-240.

[11] Barcellini, F., Détienne, F., Burkhardt, J.M., and Sack, W. 2008. A socio-cognitive analysis of online design discussions in an Open Source Software community. Interacting With Computers, 20(1), 141-165.

[12] Spada, H., Meier, A., Rummel, N., and Hauser, S. 2005. A new method to assess the quality of collaborative process in CSCL. In Proceedings of Computer Supported Collaborative Learning 2005: The Next 10 Years!, T. Koschmann, D. Suthers and T. W. Chan, Eds., Mahwah, NJ: Lawrence Erlbaum Associates. 622-631.

[13] Baker, M. 2002. Forms of cooperation in dyadic problem-solving. Revue d'Intelligence Artificielle, 16(4-5), 587-620.

[14] Dillenbourg, P. 1999. What do you mean by collaborative learning? In Collaborative learning: cognitive and computational approaches, P. Dillenbourg, Ed., Oxford: Elsevier. 1-19.

[15] Pinelle, D., Gutwin, C., and Greenberg, S. 2003. Task analysis for groupware usability evaluation: Modeling shared-workspace tasks with the nechanics of collaboration. ACM Transactions on Computer-Human Interaction, 10 (4), 281-311.

[16] Tromp, J., Steed, A., and Wilson, J. R. 2003. Systematic usability evaluation and design issues for collaborative virtual environments. Presence. Teleoperators and Virtual Environments, 12, 241-267.

[17] Hornbæk, K. 2006. Current practice in measuring usability: Challenges to usability studies and research. International Journal of Human-Computer Studies, 64 (2), 79-102.

[18] Meier, A., Spada, H., and Rummel, N. 2007. A rating scheme for assessing the quality of computer-supported collaboration processes. Computer-supported Collaborative Learning, 2(1), 63-86.

[19] Voyiatzaki, E., Meier, A., Kahrimanis, G., Rummel, N., Spada, H. and Avouris, N. 2008. Rating the quality of collaboration during networked problem solving activities. In Proceeding of 6th International Conference on Networked Learning, Halkidi. 409-416.

[20] Détienne, F., Burkhardt, J-M., Hébert, A-M., and Perron, L. 2008 Assessing the quality of collaboration in design: bridging cognitive ergonomics and CSCL approaches. Workshop « CSCW and Human Factors », CSCW'2008, (San Diego, USA, November 9 2008).

[21] Juchmes R., and Leclercq P. 2004. A Multi-Agent System for Architectural Sketches Interpretation. Eurographics Workshop on Sketch-Based Interfaces, Grenoble, France.

[22] Safin S., Boulanger C. and Leclercq P. 2005. Premières évaluations d'un Bureau Virtuel pour un processus de conception augmenté. 17ème Conférence Francophone sur l'Interaction Homme-Machine, Toulouse, France.

[23] Safin, S., Leclercq, P., and Decortis, F. 2007. Impact d'un environnement d'esquisses virtuelles et d'un modèle 3D précoce sur l'activité de conception architecturale. Revue d'interaction Homme-Machine, 8(2), 65-98.

[24] Luff, P., Heath, C., Kuzuoka, H., Hindmarsh, J., Yamazaki, K. and Oyama, S. 2003. Fractured Ecologies: Creating Environments for Collaboration. Human-Computer Interaction, 18(1), 51-84.