Attention and Engagement of Remote Team Members in Collaborative Multimedia Environments

Renate Fruchter[1] and Humberto Cavallin[2]
[1]ASCE Member, Director of Project Based Learning Laboratory (PBL Lab), Department of Civil and Environmental Engineering, Stanford University, fruchter@stanford.edu
[2]Professor, School of Architecture, University of Puerto Rico, hcavallin@uprrp.edu

ABSTRACT. An important aspect in multimedia computer mediated collaboration is to sustain the attention and engagement of remote participants during project meetings. This paper presents preliminary findings of a comparative study of two types of collaborative multimedia environments (ICT) – Webconferencing with application sharing and 3D Team Neighborhood virtual world - by evaluating the syntactic levels of micro-awareness that consist of *locus of attention* and *attention span.* These metrics provide insights into what key ICT interaction characteristics and how people attend to content presented through a collaboration interface will generate awareness, establish and sustain attention and engagement of remote participants. We used the 2009-2010 "Architecture, Engineering, construction Global Teamwork" program as the testbed. Qualitative and quantitative data was collected through field observation and EyeTracker data. Preliminary results show that meetings held in the 3D Team Neighborhood kept participants' attention 24% more time on screen vs. meetings held in Webconference with application sharing. In addition, multitasking was the typical behavior during Webconference with application sharing, whereas none or minimal multitasking was observed in the 3D Team Neighborhood.

INTRODUCTION

When collaborating in geographically distributed teams to solve problems, awareness is a crucial process in order to generate a successful interaction among the participants. In developing the awareness in computer mediated collaboration environments required for remote situations, there are multiple channels (visual, auditory) that can be used to funnel information that could cognitively engage participants in the interaction. In cases in which the interaction takes place via the use of software design for collaboration, the qualitative nature of their interfaces impacts the dynamics and level of the interaction. In this study, we explore the way in which these qualities of the interfaces affect the levels of awareness, attention and engagement during the process of collaboration in geographically distributed teams.

In order to generate and sustain the awareness of participants, there are processes that are crucial when establishing the attention and engagement of the other participants required in non-collocated collaboration problem solving. Jiazhi et al. [1] suggests that when collaborating through computer mediation, people will look at targets that help them determine whether or not their messages have been understood as intended, and that gaze patterns of speakers and listeners are closely linked to the words spoken, and help in the timing and synchronization of utterances. Vertegaal et al. [2] found that in multi-party conversations, speakers looked at the person they were talking to 77% of the time and listeners looked at the speaker 88% of the time.

According to Gutwin and Greenberg [3] , awareness has four basic characteristics:

1. Awareness is knowledge about the state of a particular environment.
2. Environments change over time, so awareness must be kept up to date.
3. People maintain their awareness by interacting with the environment.
4. Awareness is usually a secondary goal—that is, the overall goal is not simply to maintain awareness but to complete some task in the environment.

We can see that even though awareness is not a goal in itself, it is an important condition in order to achieve the proper environment for collaborative problem solving. Vertegaal et al. differentiate two different levels of awareness in cooperative work. The macro-level, refers to the awareness that conveys background information about the activities of others prior to or outside of a meeting. The micro-level of awareness according to them gives "online information about the activities of others during the meeting itself. Micro-level awareness usually has a more continuous nature than its macro level counterpart. It consists of two categories: conversational awareness and workspace awareness. Vertegaal summarizes the elements of micro-level awareness according to the attentive state, from the syntactical to the pragmatic aspects of the interaction (Table1).

Table 1. Organizing elements of awareness according to the attentive state [2]

		Attentive State	Elements	Functionality	
				Workspace Awareness	Conversational Awareness
Syntax		Locus of Attention (*Spatial*)	Location	Where are they working?	Where are the people they communicate with?
		Attention Span (*Temporal*)	Presence	Who is participating?	
			Activity	How actively are they working?	How actively are they communicating?
Semantics	Entity	Attending to Objects	Objects	What object are they using or referring to?	
		Attending to People	People	Whom do they work or communicate with?	
	Action	Attending to Actions	Action	What action are they performing or referring to?	
Pragmatics		Attention Range	Extents	What can they see?	What channels can they use?
			Abilities	What can they do?	Whom can they communicate with?
			Influence	Where can they make changes?	Where can they be?
		Future Attention	Intention (them)	What will they do next?	Whom will they communicate with next?
			Expectations (me)	What do they need me to do next?	Who wants to communicate with me next?

The syntax level contains two subcategories. The *locus of attention* describes the spatial aspects of attention, i.e., where the person directs their attention, while *attention span* describes the temporal aspects of attention, i.e., the amount of time a person can concentrate on a task without being distracted. This paper concentrates on the measurement of this particular aspect, by using methods that enable us to indirectly infer the attention of the participant, in order to establish if there is a connection between the characteristics of the interface and the way in which people attend to the interface. We are assuming in this case that the attention to the interface will be acting as an indicator of the level of engagement that the person is having in the collaboration process.

Although this variable in itself is not a unique component of the micro-level awareness in the interaction, its study provided us with important insights about how the characteristics of the interface affect this crucial aspect of non-collocated interactions.

METHODOLOGY

To explore awareness of participants in geographically distributed collaboration processes we used the "Architecture, Engineering, construction (AEC) Global Teamwork" program offered in 2009-2010 as the testbed for our study. It consisted of thirty five students engaged in the six AEC global project teams. We report preliminary findings using as examples two teams who's choice of multimedia collaborative environment was representative to the options given to all students in the program. Each of the two teams had participants at Stanford University and University of Puerto Rico. Data was collected through:

- Observations of the Stanford team members, as Dr. Fruchter was playing a dual role of mentor participant and observer during the weekly team meetings.
- EyeTracker data from a ViewPoint PC-60 Scene Camera Version eye tracker with EyeFrame hardware in the Puerto Rico site in order to identify how the *attention span* is distributed based on the gaze of the participants.

In this study gaze is a key indicator of conversational attention. As Vertegaal et. al [4] point out, results indicate that when someone is listening or speaking to individuals, there is a high probability that the person looked at is the person listened (p=88%) or spoken to (70%). According to Kendon [in 4], seeking or looking at the face of a conversational partner serves at least four functions: (1) to provide visual feedback, (2) to regulate the flow of conversation, (3) to communicate emotions and relationships; and (4) to improve concentration by restriction of visual input.

We therefore interpreted the direction of gaze as an indicator of the focus of attention of participants during meetings. Students in Puerto Rico wore the eye tracker during the meetings. For each session we recorded a total of 18 minutes of interaction. The portable EyeTracker is a non-intrusive device to track the direction of the gaze of the participants, without affecting their natural performance.

These interactions were afterwards analyzed by evaluating the position of the gaze every 12 seconds (1/5 of a minute) and correlated afterwards with the observations of the Stanford participants. The span of 12 seconds was partially decided based on information coming from literature, and also by our own observations of the meaningfulness of the interactions. Literature reports that attention in mobile interactions in laboratory conditions, as the ones we worked at for this research, takes place in an average of 12 seconds intervals [5]. We are using this parameter as a reference as there is no equivalent information connecting team interactions and span of attention to regular computer monitors.

At Stanford participants, their collaboration space, interactions, and their computer screens were observed by Dr. Fruchter in person in the PBL Lab and through her avatar in the virtual world. Instances of engagement or disengagement, side conversations, gaze foci, and use of ICT tools were noted. This provided a sense of the engagement at the individual level occurring during the team meetings. Five out of six teams had more two or three members at Stanford. Observations included: (1) How the collocated participants make their engagement (or lack thereof) visible to each other? (2) How do artifacts and ICT support or constraint engagement activities? (3) When participants engage with ICT, where is their gaze? (4) When and how did their gaze move between objects, from person to objects and back again?

TESTBED

The AEC Global Teamwork course is based on the project-based learning (PBL) methodology that focuses on problem based, project organized activities that produce a product for a client. PBL is based on re-engineered processes that bring people from multiple disciplines together. It engages faculty, practitioners, and students from different disciplines, who are geographically distributed. It is a two Quarter course that engages architecture, structural engineering, and construction management students from universities in the US, Europe and Asia.[6-7]

The core atom in this learning model is the AEC student team, which consists of an architect, one or two structural engineers, and one or two construction managers from the M.Sc. level. Each team is geographically distributed, and has a demanding owner/client that typically wants an exciting, functional and sustainable building, on budget and on time. The students have four challenges – cross-disciplinary teamwork, use of advanced collaboration technology, time management and team coordination, and multi-cultural collaboration. The building project represents the core activity in this learning environment. The project is based on a real-world building project that has been scoped to address the academic time frame and pedagogic objectives. The project specifications include: (1) building program requirements for a university building of approx. 30,000 sqft of functional spaces that include faculty and student offices, seminar rooms, small and large classrooms, and an auditorium; (2) a university site where the new building will be build, such as San Francisco, Reno, Puerto Rico. The site provides local conditions and challenges for all disciplines, such as local architecture style, climate, and environmental constraints, earthquake, wind and snow loads, flooding zones, access roads, local materials and labor costs; (3) a budget for the construction of the building, and (4) a time for construction and delivery. The project progresses from conceptual design in Winter Quarter to 3D and 4D CAD models of the building and a final report in Spring Quarter. The concept development phase deliverables of each team include: two distinct integrated AEC concepts, a decision matrix that indicates the pros and cons of the two alternatives and justifies the selection of one of the two concepts to be developed in Spring Quarter. The project development phase engages students in further iteration and refinement of the chosen alternative, detailing, modeling, simulation, cost benefit analysis and life cycle cost investigation. Spring Quarter culminates with a final AEC Team project presentation of their proposed solution, and reflection of their team dynamics evolution. The teams experience fast track project process with intermediary milestones and deliverables during which they interact with industry mentors who critique and provide constructive feedback.

All AEC teams hold weekly two hour project review sessions similar to typical building projects in the real world. During these sessions they present their concepts, explain, clarify, question these concepts, identify and solve problems, negotiate and decide on changes and next steps. Since the concepts, problems and challenges are defined by the students who work on that specific project, their level of attention and engagement is maximized. Consequently the students are highly motivated to exchange and acquire as much knowledge as they participate in the cross-disciplinary dialogue. The interaction and the dialogue between team members during project meetings evolved from presentation mode to inquiry, exploration, problem solving, and negotiation. Similar to the real world, the teams have tight deadlines, engage in design reviews, negotiate and decide on modifications. Most importantly, students learn to use and combine diverse communication channels and media to express and share their ideas and solutions. To view AEC student projects please visit the AEC Project Gallery (http://pbl.stanford.edu/AEC%20projects/projpage.htm).

MULTIMEDIA COLLABORATION ENVIRONMENTS

In this study, we compared two multimedia collaborative environments adopted by two AEC teams for their weekly meetings, in which the observed subjects were negotiating the upcoming actions. The two teams were - – Island team and Ridge team. Each team adopted a specific multimedia collaboration environment that became part of their work practice and team process. Island team adopted GoToMeeting [8] and Ridge team adopted the 3D Team Neighborhood developed in the PBL Lab at Stanford University using Teleplace [9].

GoToMeeting facilitates webconference meetings saving time and travel cost. It supports application or screen sharing. Teams can record their meeting sessions for future review. Participants can interact on the screen using the pen, highlighter tools. The organizer can change presenters and transfer control to different participants.

Teleplace provides a virtual world in which a 3D Team Neighborhood was created for each of the six AEC global project teams. It contains offices and meeting rooms, interactive rooms with multiple displays where team members can go to work, collaborate. Using avatars and unique "laser pointer" controls, participants can easily see where people are, what they are looking at, what content they are editing, and how they are using applications. In combination with Teleplace' built-in high fidelity VoIP, webcam video-conferencing and text chat, team members have an immersive environment and social cues that help them interact effectively.

EyeTracker data was collected from the two architecture students in Puerto Rico. The analysis of quantitative data from the EyeTracker was supported by in-person observations collected from four Stanford students, and 3D virtual world observations collected from the Ridge team 3D Team Neighborhood weekly meetings.

ATTENTION AND ENGAGEMENT IN COLLABORATIVE ICT SETTINGS

The following describes Island and Ridge team collaborative ICT settings according to Vertegaal's micro-level "Functionality" characteristics, i.e., workspace and conversational awareness:

- Island team was composed of an architect in Puerto Rico, a structural engineer at Stanford, an energy simulation engineer at Stanford, a construction manager at UW Madison, a life cycle financial manager at Bauhaus University, in Germany. Each of them worked in the respective university laboratory, using their laptops on WiFi, with a headset for audio. They used GoToMeeting as their multimedia collaboration environment. GoToMeeting allowed them to share their applications that were running on their individual laptops, e.g., architect showing 3D images of the building, structural engineer showing structural component options, construction manager showing cost estimates and schedules spreadsheets, life cycle financial manager showing cash flow model diagrams. GoToMeeting allows viewing, sharing, controlling one application at a time. This required participants to switch presenter and control as they were toggling between the different applications running on the different computers. It allowed all participants to view and manipulate data only on one application at a time. (Figure 1a).
- Ridge team was composed of an architect in Puerto Rico, two structural engineers at Stanford, and one construction manager in Stockholm Sweden. Each of them was working in the respective university laboratory, using their laptops on WiFi, with a headset for audio. They used the 3D Team Neighborhood in Teleplace as their multimedia collaboration environment. The 3D Team Neighborhood provided a highly immersive environment that enabled the team members to construct in real time their collaboration space around them as the dialog and interaction evolved during the meeting. Each team member could share their content on any number of displays that were created on as needed bases, as well as manipulate and annotate any content displayed in their shared workspace. All participants were able to view and interact with their content and models in context, i.e., in relation to the content and models shared by the other team members. This allowed them to interpret correlate, combine, and compare items on different displays. In addition, they were constantly aware where their team members look and are located with respect to them and the displayed content in their shared 3D Team Neighborhood. The interaction and communication in this multimedia immersive collaboration environment lead to a free and continuous flow of interaction and communication. (Figure 1b).

Based on the position of the gaze of the two architecture students at University of Puerto Rico, we used the following categories to classify the EyeTracker information:

1. *Screen*: person looking at the computer screen, specifically to the window of the application evaluated.
2. *Notes*: person taking notes related to discussion
3. *Keyboard*: person typing
4. *Outside*: person looking to anything else beyond the three previous categories.

We analyzed the videos from the EyeTracker device based on these categories, and represented the interaction graphically (Figure 1c and 1d). These preliminary results were supported by the observations made in-person at Stanford and in the 3D Team

Neighborhood showing where the students and their avatars gazed as well as their discourse dynamics. We used the notations “on task on screen” and “off task engaged-observing screen content” for the situations in which the students’ gaze was focused on the GoToMeeting or 3D Team Neighborhood content on screen. The cases in which the students were “off task and disengaged” indicated that their gaze was off screen, or on screen multitasking doing other activities unrelated to the project or ongoing discourse. Five situations were observed:

- one student on task on screen - the other(s) off task but engaged-observing content on screen or taking notes,
- one student on task on screen - the other student on a different task that is directly related to the task and topic that is discussed,
- none on task but engaged-observing screen or taking notes.
- one student on task on screen - the other(s) off task and disengaged, i.e. looking off screen or multitasking on screen,
- none on task and disengaged, (e.g., having a side conversation that is not related to task at hand or each multitasking, performing different tasks such as working on other homework, email, browsing, chatting online).

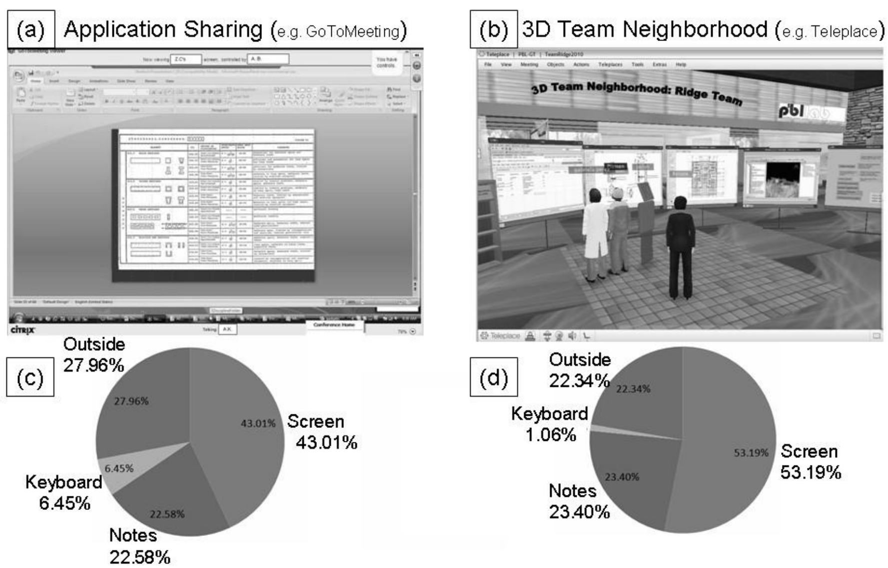

Figure 1. Preliminary Results: (a) GoToMeeting of Island Team Meeting – Structural Engineer Shares an Image of Structural Component Options to be considered by Architect, Construction Manager. (b) 3D Team Neighborhood Screenshot in Teleplace of Ridge Team Meeting Displays from Left to Right: Weekly Task List, Structural Detail Sketch, Architectural Floor Plan, Construction 4D CAD, Brainstorm Panel. (c) and (d) Distribution of time allocation by tasks during meetings in GoToMeeting and 3D Team Neighborhood.

From the preliminary qualitative observations at Stanford and in the 3D Team Neighborhoods we found that:

- The students’ and their avatars’ gaze was focused on the 3D Team Neighborhood project content on screen significantly longer periods of time and more often compared to meetings held using application sharing GoToMeeting.
- The students did almost none or minimum multitasking that was not related to the project task when running their meetings in the 3D Team Neighborhood.

However, multitasking related to different projects and other homework was the typical behavior during meetings run through GoToMeeting Webconferencing.

From the EyeTracker preliminary quantitative data analysis, it is interesting to observe that meetings held in the 3D Team Neighborhood kept participants' attention 24% more time on the screen than the in meetings held with GoToMeeting. The difference between applications is significant ($\chi^2 = 26.032$, df = 3, $p < 0.01$).

DISCUSSION

These preliminary results show that participants tend to visually engage in the 3D Team Neighborhood environment more time and more frequently than they do in the GoToMeeting with sharing environment. These preliminary observations are supported by previously reported findings in the literature, in which 3D simulated highly immersive and interactive environments seem to attract the attention and engagement of the participants in more consistent and efficient ways [10]. It is important to note that the AEC global teams had very effective team meetings once they were fluent in using the functionalities of the ICT and embedded them into their daily work practice. Nevertheless, in the case of webconferencing, participants who were not presenting or their decisions were directly impacted, tended to multitask and their gaze attention was directed elsewhere. In contrast, the 3D Team Neighborhood collaboration environment created a rich multimedia and multimodal context that kept the participants almost continuously engaged in the activity and discourse.

Attention has been studied extensively by scholars in psychology, pedagogy, neuroscience, communication and cognitive science. This study is a first step in a long term effort we started with many opportunities to extend the breadth and depth of the experiments, data collection and analysis, as well as increasing the data points.

REFERENCES

1. Jiazhi, O., et al., *Analyzing and predicting focus of attention in remote collaborative tasks*, in *Proceedings of the 7th international conference on Multimodal interfaces*. 2005, ACM: Torento, Italy.
2. Vertegaal, R., B. Velichkovsky, and G.v.d. Veer, *Catching the eye: management of joint attention in cooperative work.* SIGCHI Bulletin, 1997. **29**(4).
3. Gutwin, C. and S. Greenberg, *The importance of awareness for team cognition in distributed collaboration, Report 2001-696-19*. 2001, University of Calgary: Alberta, Canada.
4. Vertegaal, R., et al., *Eye gaze patterns in conversations: there is more to conversational agents than meets the eyes*, in *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2001, ACM: Seattle, Washington, United States.
5. Antti, O., et al., *Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI*, in *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2005, ACM: Portland, Oregon, USA.
6. Fruchter, R., *Architecture/Engineering/Construction Teamwork: A Collaborative Design and Learning Space.* ASCE Journal of Computing in Civil Engineering, 1999. **13** (4): 261-270.
7. Fruchter, R., *The Fishbowl: Degrees of Engagement in Global Teamwork.* LNAI, 2006: 241-257.
8. *GoToMeeting Webconferencing* [cited 2010 4/26]; Available from: http://www.gotomeeting.com/fec/
9. *Teleplace: Virtual Worlds Collaboration Solutions for Program Management, Virtual Operations Centers*. [cited 2010 4/26]; Available from: http://www.teleplace.com/.
10. Reeves, B. and R. Leighton, *Total Engagement: Using Games and Virtual Worlds to Change the Way People Work and Businesses Compete*. 2009, MA: Harvard Business School Publishing.