

Telepresence: Space, Self, and Stuff

Steven L. Rohall

Collaborative User Experience Group

IBM T.J. Watson Research Center, Cambridge, Massachusetts

steven_rohall @us.ibm.com

INTRODUCTION

Telepresence should be considered from three perspectives: space, self, and stuff. *Space* is the shared place in which collaboration takes place—where people become aware of others and of group artifacts. *Self* is how people are represented and how they are able to express themselves and see others within the telepresence system. *Stuff* is the artifacts of collaboration—the things that people create and manipulate. Development of future telepresence systems will be informed by issues related to these three perspectives. In this paper, I will elaborate this thesis using our experiences developing and deploying three different systems: a virtual world for supporting distributed software development teams, an e-meeting system augmented with expressive user avatars, and a collaborative slide creation tool.

SPACE: THE BLUEGRASS EXPERIENCE

In Bluegrass, we attempted to address the issue of awareness in a large, distributed software development team by creating a persistent virtual world in which the landscape was generated by team structure and populated with team artifacts. Bluegrass was embedded within Rational Jazz Team Concert, a collaborative software development environment. This virtual world featured a pastoral landscape with meeting rooms that supported simple screen sharing, awareness of people, events, and chat visualized as floating bubbles, and integration back to the host software development environment (see Figure 1). The placement of teams in the basin-like landscape was automatically computed from team hierarchy information available from the Jazz server, enabling teams to easily see other teams.

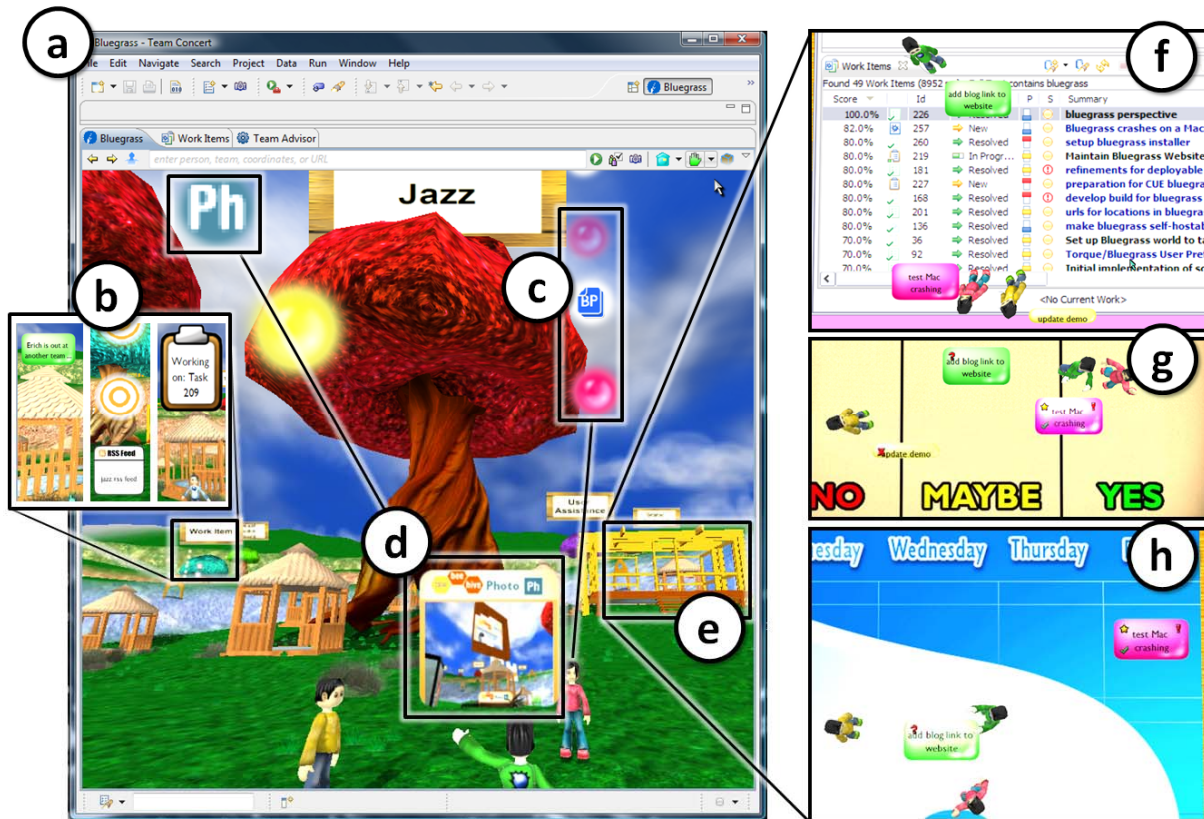


Figure 1: Bluegrass – (a) virtual world inside Rational Jazz Team Concert; (b) objects and avatars can emit bubbles viewable from afar; (c) different bubbles can be emitted, (d) such as a screenshot from a social networking profile that can be persisted; (e) team meeting areas provide top-down spaces used for: (f) application sharing and persisting chat bubbles as tasks to discuss, (g) voting and marking up bubbles with annotations like checkmarks, (h) assigning calendar tasks tied back to Jazz’s work item system

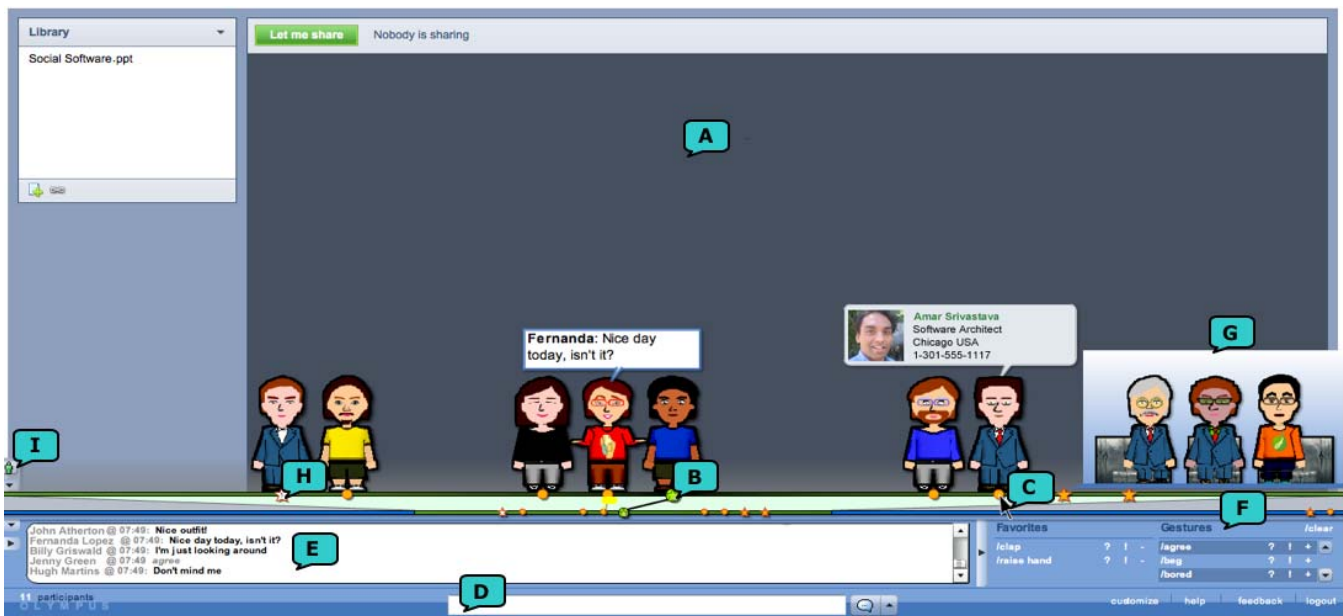


Figure 2. The Olympus interface. The main area of the interface is dedicated to slides or screen sharing (A). Every user is represented with dots along a slider (B). Hovering over a user’s dot with the mouse (C) reveals more information about that user. There is a text box at the bottom of the interface (D) that allows users to enter chat text. Chat text appears as chat bubbles as well as in a scrolling chat log (E). There is an area for users to scroll through over 30 gestures (F). The interface provides a stage for presenters and participants (G), providing meeting participants with awareness of the presenter. As indicated by a star instead of a dot, the avatar (H) is next user to go on the stage. A toggle (I) allows users to minimize the entire Olympus interface.

Observations

A virtual world seeks to immerse its users, but if those users are already immersed in work, there may be a clash along many dimensions. One of the issues we found in our Bluegrass experience is connecting the virtual space back to a real-world analogue. In Bluegrass, this was accomplished by laying out trees based on team hierarchy and connecting the virtual activities in a team’s patio deck back to the actual development environment. However, we found the virtual representation too abstract and distant from the actual work context—some found our fanciful landscape of grass, trees, and bubbles unintuitive and visually overwhelming. This raises questions for space: *Is mapping such fanciful metaphors to real work metaphors useful or distracting? When the data being presented in the space is already abstract (e.g., software code, builds, and test), how do we develop good mappings?*

Another issue centered on workflow. The way users navigated, searched, and programmed in the virtual world was too different from doing similar activities in the host Jazz environment. Adding avatars that needed to be set up and manipulated seemed like extra work. *Can we expect users to do extra work, not within their normal duties, simply to help develop awareness in a space? If not, how can their normal day-to-day activities be mapped effectively into a space?* Work spread across time zones also presents challenges for space. *If no one is in the space at the same time, what is to keep people coming back? How can we effec-*

tively represent work progress when people have come and gone?

A final issue was around work culture. While some liked the look of Bluegrass, others felt it too childish and unprofessional. And, for those who were familiar with and predisposed to use virtual worlds, the lack of avatar customizability and expressiveness were issues. Bluegrass avatars were statue-like and users did not feel connected to them. A good representation of space may not be enough—a telepresence system also needs a representation of self.

SELF: THE OLYMPUS EXPERIENCE

Based on our experiences with Bluegrass, we initiated the Olympus project with a different focus. Whereas people found the generated Bluegrass landscape difficult to map into the abstract world of software development and people were disappointed with the lack of avatar customization and expressiveness, Olympus focused entirely on avatar customization and expressiveness. Olympus was a Flash-based strip of avatars presented at the bottom of an e-meeting system that acted as a peripheral display providing awareness of meeting participants (see Figure 2). Our hypothesis was that people would take advantage of their avatar’s expressiveness to banter both before and after the scheduled meeting, much like they do in co-located meetings. To support this, Olympus had a simple avatar customization feature, similar to the Nintendo “Mii Creator.” Our avatars were cartoon-like with big heads so that expressions would be easier to see. In addition, Olympus had an elaborate gesturing mechanism. Some gestures were automatically

generated based upon user-typed chat text (e.g., the avatar would display a “dunno” gesture when the user typed a question mark). Other gestures were user-initiated (see Figure 3 for some example gestures).



Figure 3. Some avatar gestures.

Observations

People that used Olympus did enjoy creating and customizing their avatars. Within minutes, people were able to make an avatar that was quite recognizable, if they so chose, even with our very limited set of customization options. Interestingly, users customized their hair the most, followed by clothes, facial features such as beard and glasses, and head. We found that the playfulness of the avatars increased social communication among team members, possibly improving social cohesion. Some users, though, reported that the avatars got in the way and minimized Olympus so that they could maximize their view of the underlying e-meeting system (typically showing slides). This raises several questions associated with self. *How can we measure the value of a rich self-presentation like that in Olympus versus a much plainer form (like a dot or a textual name)? Although creating an avatar can be fun, do avatars really improve team cohesion or morale? Expressive avatars have demonstrated value for entertainment, but does that translate into value for the enterprise? Scalability issues aside, are avatars more useful than simple webcam video?*

People liked the avatar gesturing capability, especially automatic gestures generated by the user’s chat text. For gestures that had to be manually performed, users felt like it was an extra step compared to traditional videoconferencing. Olympus users fell into a mode where they would simply use the chat mechanism and not take advantage of the gestures unless prompted (e.g., “let’s clap”). *More broadly, how can we effectively map user emotion onto self representations in a telepresence system?*

Similarly, users were able to move their Olympus avatars left and right in a one-dimensional world that extended beyond the visible window by simply clicking. We felt that allowing users to move around would enable them to find and stand beside others they wanted to be near and with whom they wanted to converse. To prevent conversations from getting tangled and difficult to follow, we implemented a scoped chat—people could see chat messages of avatars that were visible on their screen. In reality, people moved their avatars very little in Olympus. Perhaps this was due to the limited nature of the Olympus space (although, interestingly enough, people complained when the ability to move was removed from a second iteration of the

system). More likely, this was due to the fact that moving, regardless of how simple it was to do, was perceived as extra work that was not critical to the business at hand. *Like other types of gesturing, how important is it to be able to move around in a space? Is it a critical feature of self-expression or just another chore?*

STUFF: THE LIVEDECK EXPERIENCE

Our most recent work, LiveDeck, takes a third perspective on telepresence—people are defined by their actions: what they take into the space and what they create or manipulate while they are there. LiveDeck is a system for synchronously viewing and editing slides (see Figure 4). The prototype builds upon the observation that people use slides for note taking. Slides are ideal for this purpose since they have structure (e.g., numbered lists, indentation levels) yet they also allow for non-linear thinking (e.g., text boxes can be placed wherever desired, slides are easily added, moved, and deleted). Once the meeting is over, slides also serve as a meeting record, allowing team members to share their results with those not present.

Observations

The LiveDeck prototype is still be developed, so our observations are still preliminary. However, there are several questions raised through our own use and demonstrations of the system. First is the issue of getting “stuff” into and out of the system. LiveDeck renders slides in the Open Document Format. People can seed their LiveDeck meeting with existing slides. However, LiveDeck’s rendering and editing capabilities are not as rich as stand-alone, single-user slide creation tools. As a result, people are unhappy when their slides are poorly rendered, or they are limited in the changes they can make to the slides. *How polished must tools in a telepresence system be in order to be effective?* On the one hand, people developing telepresence systems do not want to reinvent standard desktop applications that people have been using for years. On the other hand, if tools in collaborative systems are lacking, people will not want to use them. Indeed, simply taking a good, single-user application and making it collaborative (if possible), might not be successful either. Without a sense of space and self, these attempts may well fail. *What other awareness cues and collaborative features are necessary for successful collaborative editing systems?*

A second issue has to do with traditional work practices. Today, when people collaborate on slides, for example, the workflow tends to be that one person takes the lead on creating new slides and then they are forwarded to other users for comments, additional inputs, or suggestions. Some who have seen LiveDeck question whether they would ever synchronously collaborate on slides with other people. *Is this due to some fundamental truth about the way people work or is it because the available telepresence systems have only supported one style of interaction?* People know how to brainstorm around a whiteboard and people know how to pay attention when someone is presenting. Our tools today

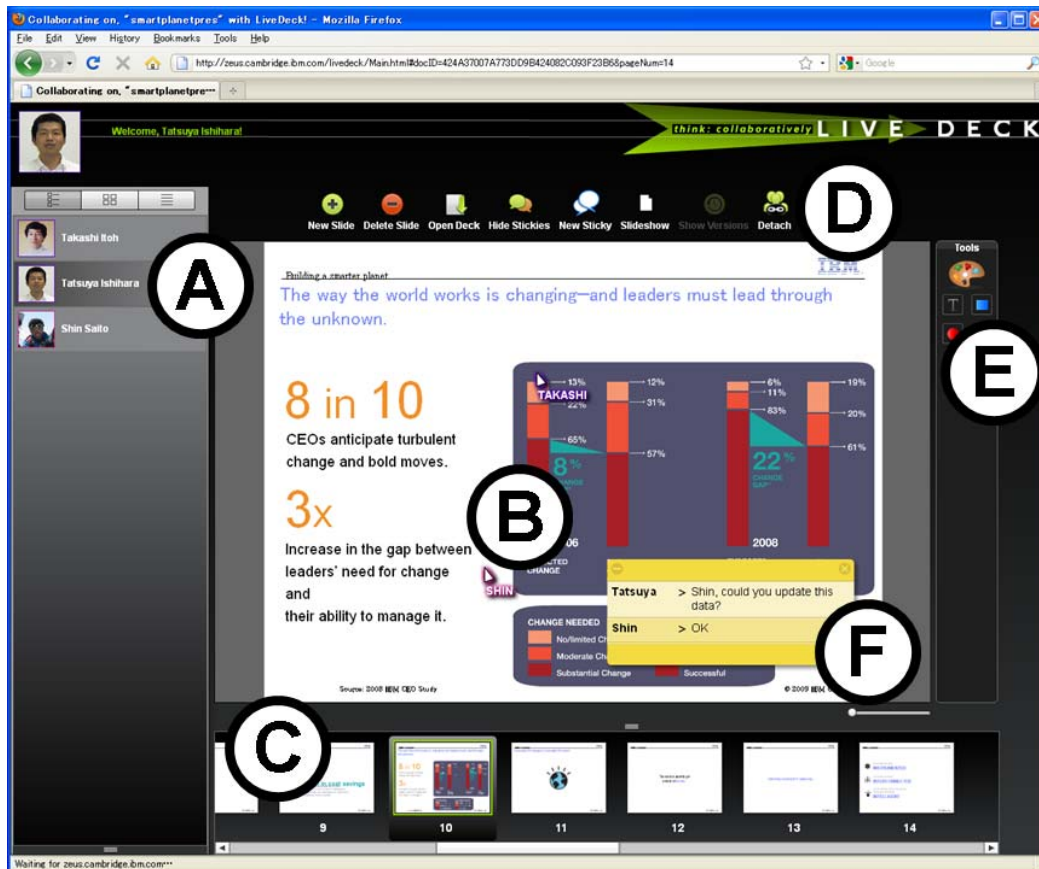


Figure 4. The LiveDeck interface. A list of users (A) shows who is currently viewing the slide deck. Telepointers of other people (B). Slide thumbnails (C) are used for navigation. In the figure, there are two highlights on thumbnail 10. The larger, gray highlight indicates the user’s current slide. The smaller, green highlight indicates the main group’s slide. By using the “detach” button (D), a user can detach from the main group and navigate the slides independently. When a user is attached to the main group, slide navigation by anyone in the group will change the slide for everyone in the group. Basic slide editing capabilities (E) that allow users to make changes. A “sticky note” capability (F) allows people to attach notes to slides.

tend to support the latter activity very well, but do not do such a good job on the former.

DISCUSSION

In this paper, I have argued that *space*, *self*, and *stuff* are three important aspects for understanding current and developing future telepresence systems. While I have used the experiences with the three systems described to illustrate the kinds of issues one will encounter, these issues are not particular to these systems. Elements of all three exist in every telepresence system. The discussion did highlight several over-arching concerns.

First, mapping abstract concepts into a telepresence space can be a problem. Mapping to real-world constructs can be useful (e.g., a conference room with table, chairs, and whiteboard is immediately recognized by people), but it may also be limiting (e.g., why should I do extra work to “walk” my avatar over to the whiteboard to use it?). Making up a new mapping can be equally dangerous. Many of today’s work processes are already abstract and every user

may have a different mental model of those processes. If the telepresence system forces people to use a particular model, users may be resistant. *How can we best map existing work practices (people, processes, documents, etc.) into a telepresence system? What are the appropriate metaphors?*

Second, while aspects of space, self, and stuff are necessary, none of them alone is sufficient for a successful telepresence system. They need to be considered holistically. *What is the right mix of space, self, and stuff for a successful telepresence system? Can certain types of teams, for example, get by without much self-presentation since the team members already know each other very well or is it exactly these sorts of teams that would benefit from fun, if frivolous, avatars?*

I look forward to developing this notion of space, self, and stuff for telepresence further from discussion with the other workshop participants.