

# Looking at the Whole Picture: A Case Study of Analyzing a Virtual Workplace

Michelle P. Steves and Emile Morse  
*National Institute of Standards and Technology*  
{msteves, emile.morse} @nist.gov

## **Abstract**

*In a previous study, we analyzed data from a field study to determine the usability problems in a groupware system. That 75% of the system use was asynchronous surprised us. Consequently, we suspected that the user-centered method that had been employed in that evaluation, might have been insufficient to detect problems related to such heavy asynchronous use. Re-analysis of the data using an artifact-centered approach revealed additional support for our initial findings and some new usability issues. In combination, we believe that user-centered and artifact-centered methods can yield superior usability analyses for systems that support both synchronous and asynchronous collaboration, and provide application developers with more appropriate priorities for addressing usability problems and system design.*

## **1. Introduction**

The current state of computing environments has changed remarkably from the days of single-user applications to one in which data, applications and users are distributed and interconnected. Increasingly, organizations are creating virtual teams that may not only be working at different locations but may well work asynchronously. At other times these same users may be collocated, working synchronously either by plan or by coincidence. To produce groupware applications that better support how workers want to work, evaluation needs to expose an application's usability problems, as well as, how workers interact with these systems and each other through these systems. Evaluation of groupware is considered difficult (e.g., [3]) and, more work is needed to better understand the problem-space.

One particular difficulty is that computer-mediated workspaces can be populated with a diverse assortment of objects, and the states of these objects are subject to change depending on a myriad of actions by the users and the system. Although groupware systems can be instrumented to capture system and user actions, it is clear that the types of interactions in collaborative environments are more

complex than in single-user applications, if only due to multiple users.

Another source of complexity in evaluating software environments stems from the large number of potential data sources. Surveys, questionnaires, interviews, video and audio recordings, system instrumentation interaction logs, and test subjects' logs are a few of the potential information sources. The completeness, quality and granularity of each of the data sources also must be considered when attempting to evaluate a system.

In a previous paper [15], we analyzed a field study that lasted for ten weeks and employed five participants. The focus of that analysis was to determine the usability of the collaboration system, and, therefore, a user-centered approach was employed. One of the surprising findings was that only 25% of the time the system was used, it was co-occupied. By using a user-centered approach, i.e., looking for each person's interactions in the data source, we overlooked an equally viable approach – an artifact-centric view of the data, i.e., reviewing how, when, and by whom artifacts were accessed or modified.

This paper reports the results of a re-analysis of the data of the study [15] from this new viewpoint. First, we describe previous, related work, and the field study. We then present the results of our artifact-centric analysis, followed by some “lessons learned” and conclusions.

## **2. Background**

The pertinent literature derives from two main areas: 1) the over-arching definition of synchrony with respect to the categorization of collaboration systems and 2) alternatives to the sole use of the user-centered approach to system evaluation.

Collaborative applications have been characterized primarily with respect to time and place. Temporally, applications may be synchronous or asynchronous. This binary distinction has been called into question in several studies (e.g., [12, 10]). Rhyne & Wolf [10] suggest that the categorization is actually harmful to the development of collaborations. They describe a model of collaboration processes, i.e., the temporal record of user actions, with

synchronous and asynchronous submodels. Sakamoto & Kuwana [12] observe that by characterizing the modes of communication by whether they happen in real-time or are temporally exclusive leads to several problems, including 1) insufficient information compatibility and 2) restriction of the system-design scope.

Whittaker et al [16] describe some “lightweight interactions” that encompass both synchronous and asynchronous interactions. Their analysis showed that there are five related, component interactions: conversational threading, one-way drop, quick connection, context preservation and regeneration, and shared object.

It remains to be seen whether relaxing or redefining the time component of collaborative systems might lead to better structured evaluation paradigms. Currently, the majority of collaboration systems that are subjected to rigorous evaluations are of one type or the other (e.g., [9]). Complexities arise when trying to evaluate systems that contain both characteristics.

When users interact in real-time, there is little need to look beyond the interactions among the participants during evaluation. However, when user interactions are asynchronous, their traces provide information that is qualitatively and quantitatively different than during synchronous periods of activity. With asynchronous work, it more difficult for the analyst to discern if the work is interdependent or independent in nature; while, synchronous work is typically interdependent. Because of this, we ask: are there alternatives to the user-centered method for evaluating these systems? Fortunately, there is some evidence in the literature about evaluating asynchronous systems. For instance, Neuwirth et al [7] discuss a task-based representation. The results of their user study suggested key capabilities required of asynchronous systems: task-specific representations, emergent representations, emergent sharing, public/private elements, incremental formalization, and asynchronous awareness. A similar technique is discussed by Olsen et al [8]. These studies lend support for an artifact-centered approach to evaluating the asynchronous aspects of a mixed-mode collaboration system.

### 3. Group work scenario

An on-going research project at NIST investigates interface standards for automated robotic welding components. In this project, several welding researchers form a geographically dispersed team working to define and test interface standards between robotic, gas-metal,

arc-welding work-cell components, controllers, and power supplies. The research is carried out using a welding testbed at the NIST facility in Gaithersburg, Maryland. New or modified equipment and interfaces can be plugged in and then tested during welding experiments. Analysis of completed welds is performed to verify effective operation of interfaces, equipment, and controllers [11]. Figure 1 shows the welding testbed, featuring the robotic arm with a welding torch and the fixturing table.



Figure 1. The remote welding testbed.

The research group consists of five to seven welding engineers and other researchers; in addition, guest researchers often participate in the work. Members of the core group are located in various buildings at the NIST site. Guest researchers are generally located off-site, occasionally in the same time zone.

The work in a welding experiment has both synchronous and asynchronous aspects. The team holds several full group meetings for planning and coordination, interspersed with periods of asynchronous individual activity, and smaller coordination meetings of two or three team members. These smaller meetings concern the “hand-off” of output from one task to other group members who need that data as input for another task, or the discussion of details that impact more than one of the experiment steps. A more detailed description of the welding scenario is given in [14].

Prior to using the groupware system, collaboration in this group was carried out using ad hoc methods. Meetings were often conducted face-to-face, with support for remote team members through telephone, email, and occasionally video- or audio-conferencing tools. Some asynchronous coordination efforts used email. Document transfer and distribution was accomplished using File Transport Protocol (FTP) and email.

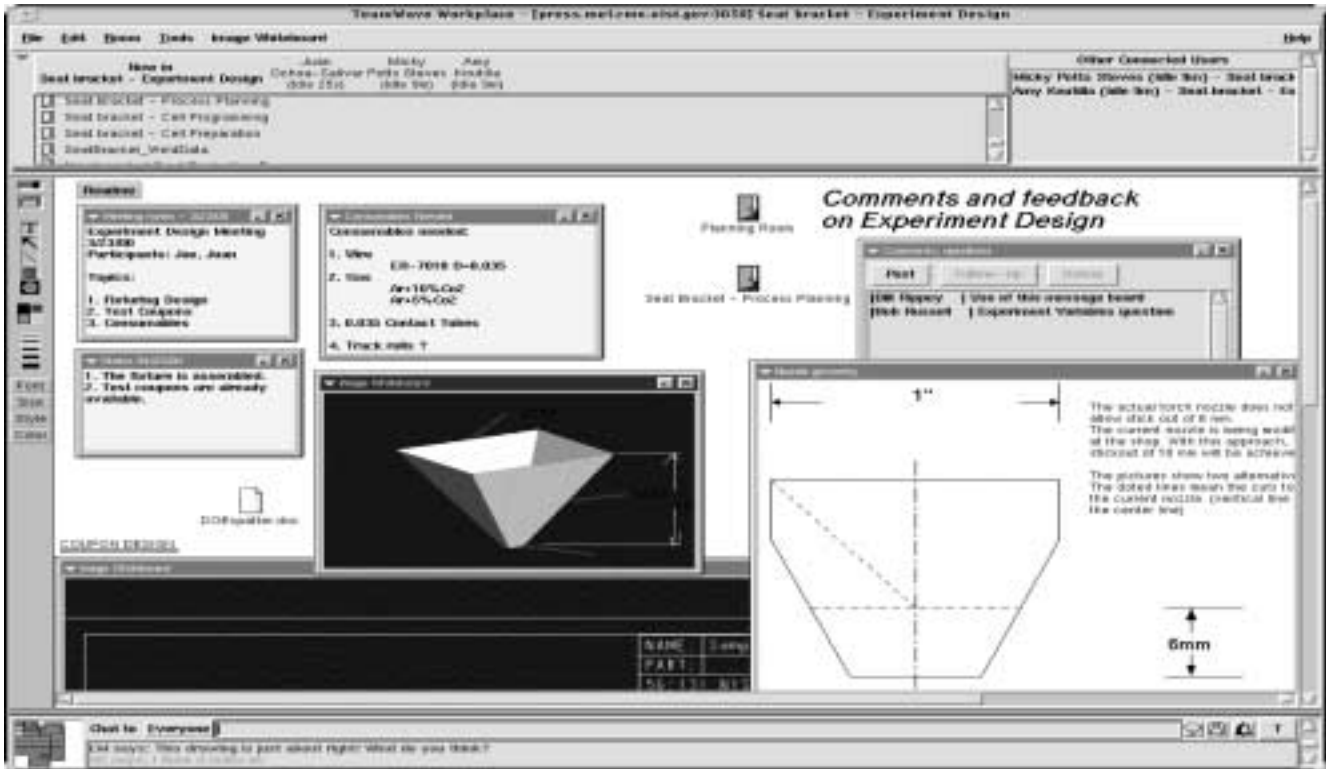


Figure 2. The Experiment Design room in TW used by the welding engineers.

#### 4. Groupware system

Teamwave Workplace<sup>1</sup> (TW) ([www.teamwave.com](http://www.teamwave.com)) was selected as a groupware tool to support the distributed welding experiments (see Figure 2). TW is a room-based collaborative system with a relaxed-WYSIWIS (What You See Is What I See) whiteboard backdrop. “Rooms” in TW provide boundaries for data groupings and user interactions, and provide a metaphor for easing the transitions between synchronous and asynchronous work [2]. Doorways provide portals to other rooms. Occupants organize data spatially within rooms by placing various tools, documents, and graphics on the whiteboard backdrop. Objects and data within the virtual space are persistent between sessions. The TW system provides for synchronous and asynchronous user interactions; but, importantly, these interactions are in the context of relevant data. Figure 2 shows the Experiment Design room in TW that was used by the welding engineers. The pull-down menus allow users to add new tools to the room and to find out about other users. The main panel is the

room's whiteboard, on which several tools and datasets have been placed. These tools with user data and user datasets are also known as *artifacts*. Participants can draw on the backdrop using drawing tools shown at left. At the bottom left of the screen is a radar overview (e.g., [13]) that shows the entire room and each person's current viewport into the room. To the right of the radar screen is a chat window where messages can be typed to some or all of the occupants of the room. To the right of the chat area are several controls for the chat tool; in particular, the “bell” button allows participants to send an alert beep to the other people in the room.

#### 5. Methodology

The study involved five steps: training, logging, questionnaires, interviews, and analysis.

**Training.** A three-hour training session and several days of individual experimentation preceded the actual work. The training session introduced the TW interface to the participants, and included hands-on use of the main features. Participants were also asked to maintain a notebook, to record any group activities in the experiment that were conducted outside of the TW environment.

**Logging and monitoring.** The welding experiment ran for 2.5 months, from the kick-off meeting to the final

1. Any commercial product identified in this document is for the purpose of describing a collaborative software environment. This identification does not imply any recommendation or endorsement by NIST.

weld. During this time, the TW evaluators monitored the collaboration, and the system logged events during individual and shared use of the tool. 7620 events were logged over the duration of the study.

**Questionnaire.** After the final welding took place, the welding researchers completed a questionnaire asking them to rate the TW interface (using a zero-to-ten scale) on several criteria. Free-form comments were also gathered.

**Interviews.** Each of the participants was interviewed to follow up survey answers, to confirm initial interpretations of the evolving data, and to address more general issues related to computer-mediated collaboration.

**Analysis.** Survey data were examined using statistical analysis to determine trends in satisfaction across the group. The system data was explored using a log visualization tool called the CollabLogger [6] and statistical methods. Additionally, the chat logs recorded by the system were analyzed to determine major topics of conversation and general communication pathways.

A summary representation was produced to tie the data together. This put all data into tabular form, with rows corresponding to the “mechanics of collaboration” [4] (a framework for groupware evaluation) and the columns holding individual data items (e.g., log results, conversation observations, survey comments). Using these evidentiary trails, we identified the major usability issues.

## **6. Analyzing the data: getting the whole picture**

As noted previously, we had collected a large volume of heterogeneous data. We had three goals: 1) to determine if the groupware system was “successful” for the welding researchers in our experiment, 2) to perform an usability evaluation, and 3) to determine if, and how, the groupware system impacted the welding work process.

The result for our first analysis goal was the easiest to obtain. From the survey and interview data, we found unanimous satisfaction by the welding researchers for the tool overall, despite some usability problems. Apparently, such a straightforward result is not typical, since “success” or “failure” can often be very difficult to discern in groupware deployment [1].

For this experiment, the usability analysis was much more involved than assessing success. As summarized in the Methodology section, we used a variety of methods to explore the datasets. The user satisfaction data helped identify potential usability issues. From these we constructed queries to examine the log data, either visually or statistically, to confirm or deny the existence of

usability problems. For example, some users reported having problems with a particular tool within TW. The log confirmed that the tool was used several times at the beginning of the experiment with no subsequent reuse – users had stopped using the tool because of its usability problem(s). Other problems were not as easy to characterize and verify as the example above; they often required broader examination of the data to pinpoint the usability problem or class of problems. The usability evaluation is documented in [15], where the top ten usability issues are identified, only one of which was related specifically to asynchronous work.

In characterizing the work performed using the groupware system, we found that participants were alone 75% of the time they were in the groupware environment. This indicated that the participants in our study were collaborating asynchronously a great deal, since they also felt that the system had basically met their collaboration needs. This made us realize that we had not explored the asynchronous-collaboration aspects of the data thoroughly. With this in mind, we re-examined the data to determine how the participants worked asynchronously, and the impact the groupware system had on their work.

### **6.1. The re-analysis: investigating asynchronous work**

We determined that there were two types of asynchronous collaboration mechanisms used in the welding experiment: email messages and TW artifacts.

**Email trails.** We had not thought to log email messages that originated outside TW, but found during the interview process that most email was created externally. During the analysis phase, we asked participants to provide us with as many of their relevant, saved messages as possible. We felt we received a good percentage of email messages sent during the experimental period, as there are no significant or logical gaps in the messages we did receive. We performed an analysis similar to the one performed for chat meetings, where major topics of conversation and general communication pathways were determined. The content gave the necessary information to determine how email messages were being used in relation to the other tools within TW.

**Artifact use.** To review how artifacts were used, we produced a summary, in tabular format: artifact name, creation date, room, creator, type of use (collaborative or independent), whether the artifact was ever used during a synchronous collaboration, and a listing of the artifact’s access/modification information. All of this information was gathered from the logged data.

## 6.2. Results of the asynchronous views into the data

Email trails. Email messages were used to communicate a variety of topics, although our participants only reported using it to coordinate synchronous meetings and notifying their colleagues of new information in an artifact or artifacts within the system. We found that email use could be characterized as shown in Table 1. It was

**Table 1: Email usage**

message categories	% of total messages
domain-specific	35.4%
meeting coordination	23.1
status/progress	23.1
artifact drop	9.2
TW-related	7.7
other technology related	1.5

surprising that the artifact drop category had such a relatively low percentage of use, since participants had reported extensive use of email for this purpose.

Artifact use. Participants used various TW tools to store and share their data, documents, and images. Types of information stored included: participant schedules and email universal resource locators (URLs), images of Computer-Aided Design (CAD) data for mark-up and discussion, experiment parameter documents and process plans, groupware space use policies (guidelines for activities within particular TW rooms), task status, meeting agendas and logs, chat logs, and open issues.

The message board tool provided an interesting middle ground between other types of tools in TW and email. For example, messages in a message board tool can be threaded like email messages. However, unlike email messages, they exist in the collaborative space, i.e., are always available. Additionally, message boards are not a “push” technology – the reader must go to space to see the message. Message board tools were used throughout the experimental period, but not nearly as heavily as email.

At the end of the experimental period, sixty-two artifacts were left in the space by the participants. Of those, all but one had been accessed by a team member other than the creator of the artifact, many multiple times as team members revisited, and periodically modified, artifacts throughout the experimental period. Additionally, nine artifacts had been used during synchronous

interactions (some artifacts were always “open” and we were not able to detect if users reviewed them during synchronous sessions unless they were mentioned explicitly). It is clear from the explicit communications and the artifact use, that artifacts housed valuable information and that the researchers used this information on an as-needed basis throughout the experimental period.

Additionally, there were thirty-nine artifacts created and deleted over the course of the experiment. Reviewing which types of tools that the welding group experimented with and then no longer used, showed the tools that either had serious usability problems or did not support this group’s work. We were able to identify those tools through the artifact-centered view, while the user-centered view did not systematically provide that insight.

Overall. Overall, the review of the asynchronous work mechanisms through an artifact-centered view revealed a fuller picture of how the team worked, how the groupware system was used, and how the project progressed over the course of the experiment. Content of both the email messages and shared objects was essential to understanding how the artifacts and email messages were used and their interrelationships.

Use of email messages and artifacts were heavily interwoven to support asynchronous collaboration. That is, email messages were used as a push technology to notify others of change or required action. Email messages often prompted participants to review materials in TW. Artifacts were used to house data thought to have team value during the welding experiment.

The additional review did reveal new usability issues; particularly, tools were identified that initially looked promising to the group, but then did not provide needed functionality. Additionally, we were better able to determine the relative importance of the usability issues from the insights gained regarding how this team worked.

While we are still assessing the impact of the groupware system on the welding work processes, it is clear from the user satisfaction data that this group believed they had improved support for their work processes with this groupware system over their previous methods, despite identified usability issues.

## 6.3. Lessons learned

Throughout our analysis, artifact and communication content was essential in understanding how tools were being used and what was happening in the work process. In an ideal data world, we would have had complete email logs and automatic, artifact versioning so that during analysis we could have been certain of having all email messages and been able to review artifact content as it changed over the course of the experiment.

Although it is effort- and time-intensive, using all the available data during analysis revealed the best picture of how the groupware system was used and the context of that use. While the user satisfaction data from the survey questionnaire and interviews was invaluable for locating usability problems, it alone did not provide a full picture of system use. We found that the participants did not always remember, appreciate, or relate the whole picture. However, the user satisfaction data, in conjunction with the log and email data, provided a rich dataset.

## 7. Conclusions and future work

In group work such as we studied, people need both synchronous and asynchronous modes of communication to support the way they work (e.g., [5]). They move between independent work and artifacts, over which they collaborate, either synchronously or asynchronously, as the task and availability of colleagues allow. Persistence of artifacts and selected communications is important and useful. Tools that support these work methods well are valuable for group work.

User-centered evaluations for groupware systems that support persistent artifacts is an incomplete evaluation perspective. Artifact-centered analysis adds another dimension of evaluation and an important perspective for understanding groupware use. Additionally, it has the potential for uncovering usability problems beyond what the user-centered approach finds, especially if user satisfaction data is not available during evaluation. Having this more holistic picture of how a system is used is important information for groupware system developers, so that they may more appropriately prioritize their efforts relative to how users of their systems work.

Logging can be implemented to provide a plethora of data that is invaluable for analysts, and is otherwise lost. Many have noted that the failure to find ways to evaluate groupware systems and learn from experience has garnered too little attention (e.g. [3]). More work needs to be focused on finding low cost methods to achieve effective levels of evaluation, while still returning a contextual perspective of a groupware system's use.

## 8. References

[1] S. Blythin, J. Hughes, S. Kristoffersen, T. Rodden, and M. Rouncefield, "Recognising 'success' and 'failure': evaluating groupware in a commercial context", Proceedings of GROUP'97 (Phoenix, AZ, Nov. 1997), ACM Press, pp. 39-46.  
[2] S. Greenberg and M. Roseman, "Using a Room Metaphor to Ease Transitions in Groupware", Research Report 98/611/02, Department of Computer Science, University of Calgary, Calgary, Alberta, Canada, 1998.

[3] S. Grudin, "Why CSCW Applications Fail: Problems in the Design and Evaluation of Organizational Interfaces", Proceedings of CSCW'88, (Portland, OR, Sep. 1998), ACM Press, pp. 85-93.  
[4] C. Gutwin and S. Greenberg, "The Mechanics of Collaboration: Developing Low Cost Usability Evaluation Methods for Shared Workspaces", Proceedings of WETICE 2000, (Gaithersburg, MD, Jun. 2000), IEEE Computer Society, pp. 98-103.  
[5] R. Harper and A. Sellen, "Collaborative Tools and the Practicalities of Professional Work at the International Monetary Fund", Proceedings of Human Factors in Computing Systems (Denver, CO, May 1995), ACM Press, pp. 122-129.  
[6] E. Morse and M. Steves, "CollabLogger: A Tool for Visualizing Groups at Work", Proceedings of WETICE 2000, (Gaithersburg, MD, Jun. 2000), IEEE Computer Society, pp. 104-109.  
[7] C. Neuwirth, J. Morris, S. Regli, R. Chandhok, and G. Wenger, "Envisioning communication: task-tailorable representations of communication in asynchronous work", Proceedings of CSCW'98, (Seattle, WA, Nov. 1998), ACM Press, pp. 265-273.  
[8] D.R. Olsen, Jr., S.E. Hudson, M. Phelps, J. Heiner, and T. Verratti, "Ubiquitous Collaboration via Surface Representations", Proceedings of CSCW'98 (Seattle WA, Nov. 1998), ACM Press, pp. 129-138.  
[9] D. Pinelle and C. Gutwin, "A Review of Groupware Evaluations", Proceedings of WETICE 2000, (Gaithersburg, MD, Jun. 2000), IEEE Computer Society, pp. 86-91.  
[10] J.R. Rhyne, and C.G. Wolf, "Tools for Supporting the Collaborative Process", Proceedings of UIST'92, (Monterey, CA, Nov. 1992), ACM Press, pp. 161 - 170.  
[11] W. Rippey and J. Falco, "The NIST Automated Arc Welding Testbed", Proceedings of the Seventh International Conference on Computer Technology in Welding (San Francisco, CA, Jul. 1997), NIST, pp. 203-210.  
[12] Y. Sakamoto and E. Kuwana, "Toward integrated support of synchronous and asynchronous communication in cooperative work: An empirical study of real group communication", Proceedings of the conference on Organizational computing systems, (Milpitas, CA, Nov. 1993), ACM Press, pp. 90 - 97.  
[13] R. Smith, R. Hixon, and B. Horan, "Supporting Flexible Roles in a Shared Space", Proceedings of CSCW'98 (Seattle, WA, Nov. 1998), ACM Press, pp. 197-206.  
[14] M. Steves and A. Knutilla, "Collaboration Technologies for Global Manufacturing", Proceedings of the ASME International Mechanical Engineering Congress and Exposition (IMECE): Symposium on Manufacturing Logistics in a Global Economy (Nashville, TN, Nov. 1999), pp. 541-555.  
[15] M. Steves, E. Morse, C. Gutwin, and S. Greenberg, "A Comparison of Usage Evaluation and Inspection Methods for Assessing Groupware Usability", Accepted to GROUP'01, (Boulder, CO, Sep. 2001).  
[16] S. Whittaker, J. Swanson, J. Kucan, and C. Sidner, "TeleNotes: Managing Lightweight Interactions in the Desktop", Transactions on Computer-Human Interaction, Vol 4, No. 2, June 1997, ACM Press, pp. 137-168.