
Groupware Evaluation: An overview

Irene-Angelica Chounta
HCI Group, Dept. of Electrical
and Computer Engineering
University of Patras, 26500 Rio
Patras, Greece
houren@upatras.gr

Nikolaos Avouris
HCI Group, Dept. of Electrical
and Computer Engineering
University of Patras, 26500 Rio
Patras, Greece
avouris@upatras.gr

Abstract

This paper discusses current trends in the field of evaluation of groupware applications. We review existing evaluation methodologies in CSCW, present the experience gained from case studies in the field and work in progress. It outlines the importance of using traditional single user evaluation methods in combination with evaluation of groupware specific issues, like communication and awareness mechanisms as well as the importance of context of use in the process.

Keywords

evaluation, groupware applications, collaboration

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

Evaluation, groupware, collaboration

Introduction

Collaborative applications (or groupware) are software systems that support people involved in a common task to achieve their goals. These systems have been developed in the broader field of Computer Supported Cooperative Work (CSCW) that describes different kinds of activities that are carried out by a group that uses information and communication technology. A similar area of study is that of Computer Supported Collaborative Learning (CSCL) that puts special emphasis on the learning outcome of the collaborative activity. The two areas share many technological and theoretical background issues. The technology in this context usually supports communication and coordination of the activity, as well as providing structure for the activity, stores record of action and intermediate and final states of the activity outcome. An activity of this nature, in general, involves two or more people who work together sharing a common objective. Evaluation of groupware applications is an open challenge for the past years and the subject of many studies. For years, research has focused on the

various aspects of collaboration that should be supported through groupware applications in order to turn collaborative activities into a success. Evaluation experts claim that until a general framework or set of guidelines is established, we should focus in evaluating the various factors of collaborative systems separately [1]. Groupware evaluation itself involves the assessment of various aspects. The objective of evaluation may be the collaborative experience, the group activity outcome, the mechanics of group interactions or communication means used, or simply the groupware application used. The objective of the evaluation study should become explicit as it is often difficult to distinguish these different aspects. For example evaluation of the technology may be done through study of a typical case of its use and often cannot be distinguished from it. One problem that has been put forward, is to define commonly acceptable methods and tools for evaluation of groupware applications. Such methods need to involve a combination of various complementary approaches that take in consideration the typical context of use of the groupware application. In the following sections we provide a brief review of the field of groupware evaluation, the experience we have gained from various evaluation studies and some ideas on the perspective of this field.

Groupware evaluation: a brief review

In general, there are some principles that groupware systems are expected to support such as, communication between partners, the establishment of shared understanding and goals, support for awareness of others' actions. In this fashion, we can define the groupware systems as the combination of technology software (applications that offer common resources,

fig. 1) and hardware devices (e.g. microphones, cameras etc) that promote collaboration among a group in order to achieve a shared goal. It is common in today's complex environments, a groupware application to contain both synchronous and asynchronous communication components, synchronous and asynchronous content sharing (e.g. file sharing or joint editors) and content management components. The scenarios of typical use of such systems are extended and complex and involve usually distributed groups of people and multiple contexts of use.

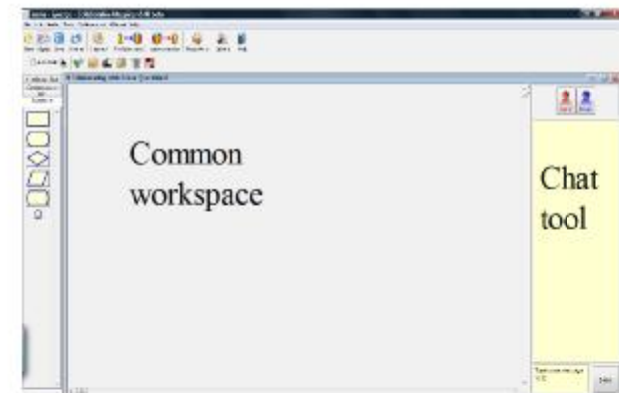


Figure 1. Example of a synchronous groupware application, it typically includes a communication tool (Chat) and a shared work space, including various awareness mechanisms.

While for single-user applications there is already an extensive corpus of knowledge relating to evaluation methods and tools, groupware evaluation methods are still in evolution. Originally some of the most well known groupware evaluation methods derive from HCI methodologies for single user evaluation as classified by McGrath in Table 1 [12]. These methods have been adapted for the collaborative context. Such methods

are: Expert based methods like use of heuristics [2,3], User testing, usually in labs [16], survey of user opinions through interviews & questionnaires, focus groups, etc, and ethnographic methods, like contextual inquiry. Among these techniques, the most prominent one involves *user studies in the lab*. For groupware systems however, it is particularly difficult to create a group in a laboratory that will reflect the real case scenarios [17].

Table 1: Evaluation Types (from [12])

		Manipulation	
		Rigorous	Minimal/None
Setting	Naturalistic	Field Experiment	Field Study, Case Study
	Controlled	Laboratory Experiment	Exploratory

Contextual inquiry in the field may be also used for groupware evaluation, however the distributed nature of such applications constrains the evaluation process necessitating high amount of resources, while it is particularly difficult to reconstruct the user experience, as this necessitates processing of large amounts of data. These restrictions led to the development of discount methods and frameworks that are based on human information processing models. We summarize some of these frameworks:

Mechanics of collaboration describe the basic operations that appear during group work towards a common goal. Collaboration, including communication and awareness aspects, is analyzed and mapped through small-scale actions described by mechanics of collaboration [8]. Frameworks such as *Collaboration Usability Analysis (CUA)* and *Groupware Heuristic Evaluation* were based on the mechanics of collaboration. Collaboration Usability Analysis (CUA) is based on a hierarchical task model for group work than involves field studies and incorporates several modified task analysis techniques. It provides high and low-level representation of collaborative activity and the interactions that take place within it [18].

Groupware Task Analysis (GTA) [20] is another analysis method and framework for the design of groupware that focuses on the triplet people, work, situation. Hierarchical task analysis in combination with human information processing models, such as keystroke level model, has also been used in groupware evaluation either in order to decompose group work to the level of individual subtasks [13] or to be integrated and adapted in collaborative settings [6].

In addition, concept-specific frameworks of study have been proposed, focusing on specific aspects of collaborative work such as awareness, communication and coordination. These frameworks put emphasis on participatory design and evaluation [19], extending data logging methods for collaborative systems [10], as well as methods for modeling and mapping awareness within a collaborative setting [14] and breakdown analysis [9].

Lessons learned from evaluation studies

In this section we briefly discuss a number of recent groupware evaluation studies that we conducted. These are indicative of the current trends in evaluation methods of this field. The presented studies are part of a more general framework that aims to provide guidelines and rules towards successful collaborative group activities. During these studies traditional HCI methods for single user application evaluation were combined with CSCW evaluation methodologies, while special emphasis was given to the effectiveness of alternative awareness mechanisms. The studies also involved qualitative analysis of video to assess the collaboration activity. We present two of these studies.

The first study [5] involved Heuristic Evaluation of a web-based argumentation tool used by communities of practice. For the evaluation of the application, heuristic evaluation for single-user interface was combined with heuristics for CSCW based on the mechanics of collaboration. Groupware Heuristic Evaluation [2] provides a set of heuristics for Shared-Workspace groupware that can be used by a number of experts for the evaluation of groupware. The groupware application supported both synchronous and asynchronous collaboration and was studied for both settings. Overall fifty (50) evaluators took part in the study. It was clear that in evaluation of groupware applications there are various aspects that need to be studied and evaluated separately but taken into consideration for an overall assessment. Expert based inspection methods should be combined with user observation methods, while we need not only focus on the collaborative functionality but also on the user interface design issues. This has been confirmed by studies that revealed that most of the issues occurring in a collaborative session are due

to inconsistencies of the interface rather than communication and awareness problems [5].

The second study [4] involved study of synchronous collaboration for problem-solving and examined allocation of attentional resources during different collaborative sessions. Three dyads' practice was monitored and analyzed. The dyads were formed in order to study different group dynamics. One dyad consisted of participants of similar knowledge background and equally motivated towards collaboration, a second dyad consisted of users of different knowledge background where the inexperienced user was monitored by the eye tracker and the third dyad was also formed by users of different background but the experienced user was monitored. The logfiles of the collaborative activity were combined with the logfiles of the eyetracker in order to analyze the interplay between task, awareness mechanisms and collaborative practice. This study revealed that awareness and communication failures are often interpreted as unwillingness towards collaboration or gradual lost of interest in a collaborative activity. The lack of adaptive awareness mechanisms that help users to set priorities towards the common goal instead of just informing about who, when and where is active, leads users to withdraw from the joint activity. They remain visible in the common workspaces, are aware of the actions of their partners but take no actual role in the collaborative activity. [4]

Another objective is to define evaluation metrics for groupware applications. Quality of collaboration has been correlated to automated metrics taken from logfiles [11]. In the same way, it is expected that data of eye movements and activity on the common

resources of groupware applications are able to provide us with useful information concerning the usability and quality of collaborative systems.

Conclusions and further research

Grudin in his article "Why CSCW applications fail: problems in the design and evaluation of organizational interfaces" [7], outlines a number of key issues in groupware design and evaluation. These concerns are still valid, despite that they have been stated over twenty years ago. Collaborative activities are difficult to model and analyze. The outcome of collaborative activities rely on many factors such as the quality of collaboration, the context of the activity and the tools that mediate the activity. The complexity of collaboration setting makes it difficult to identify the source of observed problems. So for instance tools that may support efficiently groups for playing a game may

not support students during a learning activity, or social networking tools may not be satisfactory as argumentation system. So there is still a long way to go in establishing a framework for evaluation of groupware applications, that takes all these issues in consideration and can be used in a wide range of collaborative applications. On one hand, as discussed in this paper, there is already a substantial experience in the community of research and practice and reports of good practice exist that indicate the need to use a combination of evaluation methodologies, ones relating to single-user interfaces and more specific methods relating to groupware characteristics. However an open issue remains how to include the context in which each application is used and the quality of the outcome and how to take them into account in the evaluation process.

References

- [1] Ackerman, M., Halverson, C., Erickson, T., & Kellogg, W. Resources, co-evolution and artifacts: Theory in CSCW, UK: Springer, (2008).
- [2] Baker, K., Greenberg, S., Gutwin, C. Heuristic Evaluation of Groupware based on the Mechanics of Collaboration. Lectures in Computer Science, Engineering for Human-Computer Interaction, (2001), 123-139.
- [3] Baker, K., Greenberg, S., Gutwin, C. Empirical development of a heuristic evaluation methodology for shared workspace groupware, Proceedings of the ACM conference on Computer supported cooperative work, (2002), 96 – 105.
- [4] Chounta I.A., Avouris N. Study of the effect of awareness on synchronous collaborative problem-solving. In 2nd International Conference on Intelligent

Networking and Collaborative systems Proceedings – INCoS, (2010), 153-160.

- [5] Chounta, I.A., Avouris, N. Heuristic Evaluation of an Argumentation Tool used by Communities of Practice. In EC-TEL Workshop: TEL-CoPs'09: 3rd International Workshop on Building Technology Enhanced Learning solutions for Communities of Practice, (2009).

- [6] Ferreira, A., Antunes, P. A Technique for Evaluating Shared Workspaces Efficiency, in Computer Supported Cooperative Work in Design III, Lecture Notes in Computer Science, 4402, Springer, (2007), 82-91.

- [7] Grudin, J., Why groupware applications fail: Problems in design and evaluation, Office: Technology and People, 4(3), (1988), 245-264.

- [8] Gutwin, C., Greenberg, S. The Mechanics of Collaboration: Developing Low Cost Usability Evaluation Methods for Shared Workspaces, IEEE 9th International

Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WET ICE'00), (2000), 98.

[9] Hartswood, M., Procter, R. Design guidelines for dealing with breakdowns and repairs in collaborative work settings. In *International Journal of Human-Computer Studies*, 53(1), (2000), 91-120.

[10] Helms, J., Neale, D.C., Isenhour, P.L. and Carroll, J.M. Data Logging: Higher-Level Capturing and Multi-Level Abstracting of User Activities. In *Proceedings of the 40th annual meeting of the Human Factors and Ergonomics Society*, (2000).

[11] Kahrimanis, G., Chounta, I.A., Avouris, N. Study of correlations between logfile-based metrics of interaction and the quality of synchronous collaboration, 9th International Conference on the Design of Cooperative Systems, Workshop on Analysing the quality of collaboration, (2010). Published in *International Reports on Socio-Informatics (IRSI)*, 7(1), 2010

[12] McGrath, J.E. Methodology matters: doing research in the behavioral and social sciences. In *Human-computer interaction*, Ronald M. Baecker, Jonathan Grudin, William A. S. Buxton, and Saul Greenberg (Eds.). Morgan Kaufmann Publishers Inc., 1995, 152-169.

[13] Min, D., Koo, S., Chung, Y.H., Kim, B. Distributed GOMS: an extension of GOMS to group task. In *Systems, Man, and Cybernetics, IEEE SMC '99 Conference Proceedings*, 5, (1999), 720-725.

[14] Neale, D.C., Carroll, J.M, Rosson, M.B. Evaluating computer-supported cooperative work: models and

frameworks. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work (CSCW '04)*. ACM, (2004), 112-121.

[15] Olson, G. M., and Olson, J. S. Technology support for collaborative workgroups. In G. M. Olson, J. Smith, and T. Malone (Eds.), *Coordination theory and collaboration technology*. Hillsdale, NJ: Lawrence Erlbaum Associates, (2001), 559-584.

[16] Olson, G. M., and Olson, J. S. User centered design of collaboration technology. *Journal of Organizational Computing*, 1, (1991), 61-83.

[17] Pinelle,D., Gutwin, C. A review of groupware evaluations. WETICE '00: 9th International Workshop on Enabling Technologies, IEEE Computer Society, (2000), 86–91.

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

[19] Ross, S., Ramage, M., Rogers, Y. PETRA: participatory evaluation through redesign and analysis. In *Interacting with Computers*, 7(4), (1995), 335-360.

[20] Van der Veer, G.C, Lenting, B.,F., Bergevoet, B.A.J. GTA: Groupware task analysis -- Modeling complexity, *Acta Psychologica*, 91(3), Usage of Modern Technology by Experts and Non-professionals, (1996), 297-322.