Heuristic Evaluation of an Argumentation Tool used by Communities of Practice.

Irene-Angelica Chounta, Nikolaos Avouris,

University of Patras - Human-Computer Interaction Group, Electrical and Computer Engineering Department, GR-265 00 Rio-Patras, Greece
{houren@ece., avouris@}upatras.gr

Abstract. This paper reports a systematic evaluation study of a groupware application that was build to support communities of practice. The purpose of this study was to apply existing Evaluation Methodologies on groupware applications, record the results and point out the weak points of the evaluation scheme. Our main objective is not to evaluate the particular application but to record the whole evaluation process that was followed in order to reflect on the characteristics of the approach, point out the difficulties and significant parts of our study for further research in the field of evaluation of communities of practice groupware tools.

Keywords: heuristic evaluation, CSCW, communities of practice.

1 Introduction

Communities of Practice have existed since early history of society, however the term in its current use was first introduced by Lave and Wenger [10], [11] who examined communities of practice as settings for learning. They introduced their theory of learning through a process called *Legitimate Peripheral Participation* which relates to apprenticeship, i.e. the process through which newcomers to the community learn from experts, as they are allowed to undertake more and more tasks and gradually move to full participation. As a result of these theoretical concerns, study of communities of practice has evolved in recent years, improving our understanding of learning as an act of participation and therefore as a social activity. The main characteristics of such setting include the formation of a community, the members of which are involved in some kind of collaborative learning and in which co-exist members of disperse skills and expertise, so that apprenticeship processes can take place. Support of communities of practice by tools is also an active area of research and practice. Community knowledge repositories are such typical tools as knowledge is integral with the life of communities of practice, while groupware applications that support interaction between community members are also needed, in order to structure community activities, maintain repository of interactions and facilitate remote communication.

In this context an important concern is related to systematic and effective evaluation of such tools. An area of research and practice that can be proved particularly useful is that of groupware evaluation methods. Already many researchers have expressed their views and made ambitious proposals for new methodologies [1], [9], while the complexity of the field has been acknowledged, since it combines usability, individual psychology, group dynamics, efficiency of awareness and communications, organizational and cultural structures and so on. The so far proposed methods mainly focus on single aspects such as usability, awareness and communication efficiency. In this way the results were rather difficult to use and the whole field was characterized by "the gap between social requirements and technical feasibility" as Marc Ackerman has noted [2].

This paper attempts to contribute to the discussion on systematic evaluation of tools to support communities of practice, by reporting the method and the main findings of a study that had as an objective to evaluate a prototype that was designed to support argumentation between members of communities of practice. The paper is divided into the following sections: first a description of the Application that was evaluated is provided, then a brief introduction of the experimental setup, followed by the main findings of the two phases of the study, i.e. heuristic evaluation of the user interface and heuristic evaluation of Groupware functionality based on the Mechanics of Collaboration. It is hoped that the reported study may be useful for practitioners and researchers that are in search of systematic methods for design and evaluation of tools supporting communities of practice.

2 Description of the application and experimental setup.

2.1 Description of the application.

The application that was evaluated, is a web-based tool supporting argumentative collaboration towards learning and among members of Communities of Practice (CoPs)[12], [13]. It provides a common workspace for discussion among CoPs participants, makes use of various visualizations and reasoning mechanisms. Users can create new communities or join existing ones and collaborate on various aspects. Emphasis has been given on the spatial organization of a discussion and in the way that conclusions may arise from such an organization.

We used this application to support two groups of users of mixed expertise, as discussed in section 2. One group was consisted of usability evaluators who took part in a guided, heuristic evaluation. The other group was consisted of designers who had to design a tool for remote evaluation of services and products. Since the users of both groups were of mixed expertise, they were adopting roles of apprentices and tutors. By tracking arguments, peer learning emerged. Communities of practice was a term who attempted to explain and describe learning that occurs in apprenticeship situations [10], [11]. In that sense we can describe both groups as communities of practice.

2.2 Experimental setup.

Two evaluation studies of the same web-based groupware system were conducted and are reported in the rest of the paper. The system that was evaluated was designed to support communities of practice. In both studies the system was used by small groups of peers of varying levels of expertise, as a tool for collaboration and building of group knowledge in relation to given tasks. Peer learning and development of skills of design and evaluation of interactive systems was the main objective of these activities. This is a task in which one expects that group members learn by doing and participating in the discussion, while less experienced group members play the role of apprentices of the community. The argumentation tool in this context acts as a repository of group knowledge and facilitator of interaction of peers that are located remotely.

In the first study, we used the system in order to support a group of final year undergraduate and postgraduate students of interaction design in the Electrical and Computer Engineering Department of the University of Patras. The purpose of the activity was to conduct a study using an inspection method called heuristic evaluation. The students' task was *to evaluate the user interface of a groupware system, using a given set of heuristics*. In order to do that they had to collaborate with the other members of their team using the system, exchange ideas and state their opinion. The allocated time for completing the task was a week. Thirty nine (39) students participated in this study, divided in four groups: three groups of ten members each and one group of nine. The main means of interaction was the argumentation tool, which was used for building a common view of the usability problems found according to the used heuristic rules. Due to the relatively long duration of the experiment, group members interaction was mostly asynchronous. After completing the experiment, the participants filled out a questionaire in order to evaluate the usability of the application (heuristic evaluation) and their experience while using it.

In the second study, the activity of a focus group was supported by the argumentation tool. The task of the group was *to discuss and identify the key requirements for the design of a new tool*[1]. The communication between the members of the focus group was carried out using the web-based argumentation support system that was under evaluation. The group members were ten designers of varying degree of expertise, they included web design practitioners, postgraduate and undergraduate students of various related disciplines (e.g. computer science, design, usability etc.). This community was formed during a workshop on usability evaluation. The duration of the specific task was thirty five (35) minutes, so there was need for synchronous communication. The task was structured, so during the allowed time, the participants had to discuss and answer three related questions. At the end of the procedure the participants were interviewed about their experience using the application.

For both the experiments log files concerning the interaction with the application were collected.

1 The web tool that was to be designed supports remote evaluation studies.

3 First study: Heuristic evaluation of the argumentation tool.

Heuristic evaluation is considered one of the most widely used expert based inspection methods for evaluation of interactive systems. One common way of conducting this evaluation is to examine applicability of some heuristic principles on a given interactive system. A widely used set of heuristics has been propose by Nielsen [6], [7] for general purpose interactive systems. These principles, shown in Table 1, are called "heuristics" because they are more in the nature of rules of thumb than specific usability guidelines.

Table 1. Heuristic Principles applied in first study.

1. Visibility of system status
2. Match between system and the real world
3. User control and freedom
4. Consistency and standards
5. Error prevention
6. Recognition rather than recall
7. Flexibility and efficiency of use
8. Aesthetic and minimalist design
9. Help users recognize, diagnose, and recover from errors
10. Help and documentation

The usual way to conduct heuristic evaluation is by involving independent experts. 5-6 experts are considered enough for finding most of usability problems. In our case we used a variation of this approach: Due to the nature of the system under evaluation (web-based application), we used a collaborative heuristic evaluation approach that is more suitable for communities of practice. We asked the 'experts' to work collaboratively and use the tool that they were supposed to evaluate in order to discover the faults in the design, using the heuristics. This was in order to make the experts have hands on experience of use of the argumentation tool, before they proceed to its evaluation. The outcome of the study is manifold. On one hand from the records of the tool, objective data can be obtained. On the other hand the evaluation report produced by the teams of 'experts', gives a structured record of their view on the limitations of the argumentation tool. It should be noted that this phase of the evaluation is based on general purpose heuristic principles that are suggested for single user applications. In the next section more specific heuristics will be applied, related to collaborative and shared-space tools.

Table 2. Comments of phase #1 heuristic evaluation of argumentation tool.

User comments on heuristics violated	#users
The options "menu" and overview tool "minimap" provided by the application are not stated as clickable, though they are.	8
The pop-up windows of the application use the symbol "x" on top right corner to state how to close a window. Instead of this the windows of menu, mini map and broadcast use a "hide" button placed in different spots.	14

The interface does not support the actions of undo and redo.	34
It is not clear to the user how to edit or delete an existing object or even if these actions are supported.	32
The interface does not make use of shortcuts for most used actions or keyboard functions for usual activities such as copy and paste.	25
The objects added or modified since the last view of the workspace are not highlighted in any way resulting in difficulty of viewing the workspace.	17
The language used by the application is often not understandable by the users.	17
It is not clear that in order for a user to talk or work in the shared space should first make a request and wait for his turn	22
The users could not easily find how to chat with other participants.	14
Users were confused concerning the privacy of their chat messages.	5
The images used to represent the objects in the workspace were rather large and confusing.	5
The information button was not providing helpful information to the users as expected.	5
Chat does not provide feedback whether a message is composed or send	7
The function of search that the application provides does not work as expected and confused the users on how they are supposed to perform a routine search.	11
The participants are not presented in the workspace as expected. The application makes use of large images that are space consuming and confusing.	8
Comments inserted in the workspace are not re-sizable and therefore not easy to view.	14
The system does not make use of confirmation messages so as to prevent actions that take place by mistake.	22
The system does not provide enough mechanisms of feedback so as to inform the user of its current state.	6

Table 3. Application's interface evaluation Results according to Nielsen's Heuristics.

	Support		
Heuristics.	Poor	Satisfying	Full
1. Visibility of system status	x		
2. Match between system and the real world		x	
3. User control and freedom	x		
4. Consistency and standards		x	
5. Error prevention	x		
6. Recognition rather than recall	x		
7. Flexibility and efficiency of use	x		
8. Aesthetic and minimalist design		x	
9. Help users recognize, diagnose and recover from errors	x		
10. Help and documentation.	x		

Based on the results of this first experiment where 'experts' were asked to conduct a Heuristic Evaluation Study of the requested application, Table 3 was built. In detail, each user created a report that presented his evaluation study. Each report included a table of two columns. One column that describes the problems of the application that have been identified by the user and a second column of the heuristics that have been violated respectively to the problem reported in the first column. An expert usability evaluator reviewed the students' reports and summarized the comments on heuristic violations and problems identified. Based on these reports, the usability expert created Table 3.

According to this table, users found that seven out of the ten heuristic principles were violated while the other three were supported in a satisfying degree but not fully.

4 Second Study: Heuristic Evaluation of the argumentation tool based on Mechanics of Collaboration

As discussed in section 3, general purpose heuristics may be applied in evaluation of a groupware application. However some more specific principles are needed for evaluation of the functionality related to collaborative aspects of the application. Baker, Greenberg and Gutwin proposed an evaluation method for groupware applications that applies specific heuristics for such systems [3], [4], [5].
According to this methodology, there are eight heuristics, called *Mechanics of Collaboration*, that need to be taken in consideration when evaluating a groupware system. These are included in Table 4.

In this section we use the data, gathered from our experiments, and applied these heuristics. The data included interviews from the users, screen recordings and log files.

Table 4. Heuristics for evaluation of groupware: the mechanics of collaboration.

1. Provide the means for Intentional and Appropriate Verbal Communication.
2. Provide the means for Intentional and Appropriate Gesture Communication.
3. Provide Consequential Communication of an individual's embodiment.
4. Provide Consequential Communication of Shared Artifacts.
5. Provide Protection.
6. Measurement of Tightly and Loosely-Coupled collaboration.
7. Allow people to coordinate their actions.
8. Facilitate finding collaborators and establishing contacts.

Heuristic1. Provide the means for Intentional and Appropriate Verbal Communication.

The application that we evaluate does not provide digital audio link for direct verbal communication, but as most groupware applications, provides a text chat facility for participants. The chat tool is implemented via parcel-post (e.g. type a line and send) and the chat messages are broadcasted to all the members of the workspace, as well as recorded for later viewing.

Text is useful for short interaction and commonly used, though it provides limited communication. Therefore it is very important that the tool supports with no flaws the chat activity. The chat tool of the application, as we can see in Fig.1, is relatively small comparing to the size of the workspace, even though it can be hidden. This makes reading and writing for a user quite difficult. Moreover there is no mechanism to notify the users – especially when the chat window is not expanded- that there is broadcasted chat activity. This has as a result users to ignore chat messages frequently.

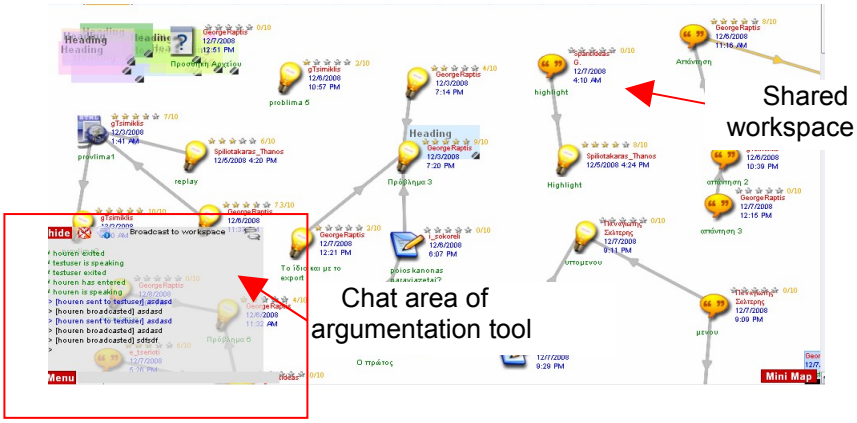

Fig. 1. *Workspace and chat area of the application.*

Heuristic 2. Provide the means for Intentional and Appropriate Gesture Communication.

Users of the application cannot work contemporarily in the shared workspace. However the system provides a kind of telepointer. When a user clicks the alarm mechanism the pointed area in the workspace is highlighted for all the participants.

Heuristic 3. Provide Consequential Communication of an individual's embodiment.

This heuristic is not supported at all by the application.

Heuristic 4. Provide Consequential Communication of Shared Artifacts.

When an action in the workspace is completed, and the producer of the action has finished working in the workspace, the artifact inserted or edited becomes visible for the group, while the name of its producer is shown, see fig.2. However the application does not make use of any sounds for declaring objects nature or state.

Fig. 2. *Object representation in the workspace.*

Heuristic 5. Provide Protection.

The application implements turn taking in order to prevent inconsistencies and conflicts between users' simultaneous actions. Participants have to make a request in order to use the shared workspace and they can only do so if they are granted permission. While the space is edited, other participants cannot view the changes made until their turn comes. Permission to talk is given to users according to the time they made the request and apart from that, the application supports the role of the coordinator who can deny a request or change the turn of a community member.

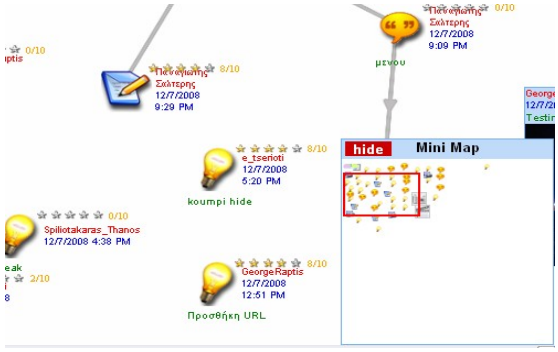

Fig. 3. *Minimap facility*

Heuristic 6. Measurement of Tightly and Loosely-Coupled collaboration.

The system provides no viewports or overview panels except from a minimap, see fig.3, and a replay mechanism. The minimap displays the whole workspace in small size but it cannot be resizable for more detailed viewing. In that sense, a participant who needs to focus on another part of the workspace is not able to follow with the rest of the activity. The replay mechanism allows a user to playback the whole activity up to that moment but while replaying, viewing of current workspace is prevented. However, the turn taking mechanism that the application makes use of, does not allow working in parallel or viewing workspace anyway.

Heuristic 7. Allow people to coordinate their actions.

The coordination of users' action is made by the turn taking mechanism or by the coordinator of the activity. For this reason the participants are not fully aware of others actions within the common workspace as well as they are not allowed to coordinate their interaction as in real world.

Heuristic 8. Facilitate finding collaborators and establishing contacts.

Participants are able to view which users are online and actively working on a specific workspace. They are also able to send email or chat messages to other users.

From Table 5, we view that support for heuristics for verbal communication, consequential communication of shared artifacts and facilitation of finding collaborators and establishing contacts is adequate according to the users who participated in the two experiments. Moreover, the turn taking mechanism provides full protection against inconsistencies in the workspace from simultaneous actions. However due to this mechanism the application does not provide freedom for users to coordinate their actions and management of tightly and loosely-coupled collaboration.

Table 5. Heuristic Evaluation based on the Mechanics of Collaboration.

	Support				
Heuristics	Not Existing	Poor	Average	Satisfying	Full
1. Means for Verbal communication			x		
2. Means for Gesture communication		x			
3. Consequential Communication of individual's embodiment	x				
4. Consequential Communication of Shared Artifacts			x		
5. Protection					x
6. Management of Tightly and Loosely-Coupled Collaboration	x				
7. Freedom of Coordinating actions	x				
8. Facilitation of finding collaborators and establishing contacts.			x		

5 Conclusions.

In this study we attempted to qualitatively evaluate an argumentation support, groupware application used by Communities of Practice The evaluation study was divided in two parts: evaluation of the user interface of the application according to general purpose heuristic principles and evaluation of the collaborative aspect of the application using the Mechanics of Collaboration. An experimental design was presented that included two communities that worked on more lose and more tight time constraints, thus necessitating asynchronous and synchronous communication respectively. The characteristics of the groups (i.e. varying degrees of expertise of the group members) and the tasks (collaborative evaluation and collaborative design) had the characteristics of communities of practice and peer learning. Based on the rich set of collected data we record the heuristics violated, the circumstances under which the violations have taken place and the frequency of the violations. Both objective (observations) and subjective (experts' opinions) were used in the processAs a final conclusion, from the reported experience, we should deduce that for evaluation of groupware applications, like the argumentation tool discussed here, there are various aspects that need to be studied. Expert based inspection methods, like heuristic evaluation need to be combined with user observation methods, while we need not only focus on the collaborative functionality but also on the single user interface design characteristics. From our study we realized that most of the issues that were

found during the use of such a system are due to inconsistencies during the design of the interface rather than communication and awareness problems.

References

1. Grudin,J. Why CSCW Applications Fail:Problems in the design and evaluation of organizational interfaces. Proceedings of the 1988 ACM conference on Computer-supported cooperative work, 85- 93 (1988).
2. Ackerman, Mark S.: The Intellectual Challenge of CSCW: The Gap Between Social Requirements and Technical Feasibility. Human-Computer Interaction,15:2,179-203 (2000).
3. Gutwin,C., Greenberg,S.: The mechanics of collaboration: Developing low cost usability evaluation methods for shared workspaces. 9th IEEE International Workshops on enabling technologies, (2000).
4. 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, 123-139 (2001)
5. Baker, K., Greenberg, S., Gutwin,C.: Empirical development of a heuristic evaluation methodology for shared workspace groupware, Proceedings of the 2002 ACM conference on Computer supported cooperative work, 96 - 105 (2002)
6. Nielsen,J.: Usability Engineering, Morgan Kaufmann Publishers Inc., San Francisco, CA (1993).
7. Nielsen,J., Mack,R.L.: Usability inspection methods, John Wiley & Sons, Inc., New York, NY (1994)
8. Ackerman, M., Halverson, C., Erickson, T., & Kellogg, W.: Resources, co-evolution and artifacts: Theory in CSCW. London, UK: Springer (2008)
9. Ramage, M.: The Learning Way: Evaluating Cooperative Systems. Unpublished PhD thesis, University of Lancaster. Available at http://www.comp.lancs.ac.uk/computing/research/cseg/projects/evaluation/index.html (1999)
10. Lave,J., Wenger,E.: Situated Learning: Legitimate Peripheral Participation, Cambridge University Press(1991)
11. Lave,J., Wenger,E.: Communities of Practice: Learning, Meaning, and Identity, Cambridge University Press (1998)
12. Tzagarakis,M., Karousos,N., Karacapilidis,N.: On the Development of Web-based Argumentative Collaboration Support Systems. In Proceedings of the 2nd World Summit on the Knowledge Society (WSKS 2009), Springer-Verlag, Berlin, Lecture Notes in AI (to appear) (2009).
13. Karacapilidis,N., Tzagarakis,M.: Supporting Argumentative Collaboration in Communities of Practice: The CoPe_it! approach. In: N. Karacapilidis (ed.), Solutions and Innovations in Web-Based Technologies for Augmented Learning: Improved Platforms, Tools and Applications, IGI Global, Hershey, PA, USA, pp. 245-257 (2009).