

Heuristic evaluation for mobile groupware: Evaluating two Audience Response Systems

Luciana Pereira de Araújo

Universidade do Estado de Santa Catarina. Rua Paulo Malschitzki, S/N, 89219-710, Joinville, SC, Brazil
luciana.pda@gmail.com

Marcio José Mantau

Universidade do Estado de Santa Catarina. Rua Dr. Getúlio Vargas, 2822, 89140-000, Ibirama, SC, Brazil
marcio.mantau@gmail.com

Jucilane Rosa Citadin, Carla Diacui Medeiros Berkenbrock, Avanilde Kemczinski

Universidade do Estado de Santa Catarina. Rua Paulo Malschitzki, S/N, 89219-710, Joinville, SC, Brazil
jucilane.rosa@gmail.com, diacui@joinville.udesc.br, avanilde@joinville.udesc.br

Gian Ricardo Berkenbrock

Universidade Federal de Santa Catarina. Rua Presidente Prudente de Moraes, 406, 89218-000, Joinville, SC, Brazil
gian.rb@ufsc.br

Mauro Marcelo Mattos

Fundação Universidade Regional de Blumenau. Rua Antônio da Veiga, 140, 89012-900, Blumenau, SC, Brazil
mattos@furb.br

Abstract. In conferences, congresses and symposia, there are issues that may be handled by applications running in mobile devices. In these environments, Audience Response Systems can be used to increase iteration between participants. This paper aims to evaluate the usability of two Audience Response Systems, namely Simple Question and Voting, developed for collaborative proposals. We employed a set of 15 heuristics for mobile groupware evaluation, which includes three fundamental aspects of mobile groupware applications: HCI, mobility, and collaboration. In the evaluation performed in this work, three specialists used the heuristics to evaluate both systems interfaces. The evaluation was conducted in three steps: (i) exploration of the system and its features; (ii) usability evaluation; (iii) consolidation of the collected data and reflection. The evaluation results of Simple Question and Voting indicated that both are in accordance with some usability aspects. However, we also identified common problems, such as non-appropriate communication methods, low coordination, and group management. Based on the results and experiences obtained over this evaluation, we argue that the heuristic set presented is applicable to evaluate mobile groupware usability.

Keywords: heuristic evaluation, mobile groupware, usability.

Introduction

The use of mobile devices during lectures enhances the interaction between audience and speaker (Teevan *et al.*, 2012). Audience Response Systems, Student Response Sys-

tems, Classroom Student Systems, and Clickers are terms used to define systems used during presentations in order to provide feedback to the speaker, as well as increase interaction among participants (Rechenthin and Molenda, 2009).

Audience Response Systems allow rapid searches, anonymously and the results can be immediately used during the lecture. According to Esponda (2008) and Scornavacca *et al.* (2009), among the benefits of using Audience Response Systems in classrooms one can cite: increased interest, participation, understanding and discussion among students, improved teacher awareness of students difficulties; promoting more effective learning; providing feedback to the teacher; enabling increased student motivation, and enabling collective learning. Rechenthin and Molenda (2009) show that these systems can increase student engagement in the classroom.

This paper aims to evaluate the usability of two Audience Response Systems developed for collaborative proposals: Simple Question, and Voting. These groupwares were chosen because they are new groupwares that have not been evaluated yet.

The evaluation of collaborative systems is an important issue in the collaborative field (Antunes *et al.*, 2012) but not yet fully solved (Herskovic *et al.*, 2007). There are several specific evaluation methods for mobile groupware that have been proposed, as presented by Herskovic *et al.* (2007) and Antunes *et al.* (2012). According to Pinelle and Gutwin (2000), an interesting category of groupware evaluation is inspection, in which experts analyze the system interface searching problems that may hinder the system use (e.g. usability inspections). These inspections are generally based on a checklist of guidelines or desirable characteristics [heuristics] that a system should satisfy (Antunes *et al.*, 2012).

The Heuristics Evaluation (HE) method is considered faster and cheaper than other evaluating methods (e.g. experiments). Regarding the mobile groupware context, these inspection methods are called Groupware Heuristics Evaluation (GHE). Antunes *et al.* (2012) say that GHE is concerned with the effectiveness, efficiency, precision and the user's satisfaction.

The main goal of this paper is to employ a set of heuristics based on Baker's and Bertini's heuristics to make evaluating the usability and collaboration of mobile groupwares possible. This adjusted set of heuristics will be used to evaluate two Audience Response Systems.

This article is organized as follows. The section "Related Works" presents the related work with this paper. The section "Audience response system" describes the Audience Response Systems evaluated in this work: Simple

question and Voting. The section "Heuristics for mobile groupware evaluation" presents a set of heuristics selected for mobile groupware evaluation. The section "Evaluation methodology" presents the evaluation methodology and the results of the heuristic evaluation. Finally, we describe the final remarks and conclusions.

Related work

In this section we present Bertini's and Baker's heuristics. These heuristics were used to evaluate the usability and collaboration of two audience response systems. Next, we present some works related to audience response systems.

There are some Heuristics Evaluations (HE) and Groupware Heuristics Evaluations (GHE) proposed in the literature, however, they do not usually include aspects that are essential for mobile groupwares. For example, Nielsen's Heuristics (Nielsen and Mack, 1994) address only usability aspects desirable for conventional interfaces. This set does not address issues related to the context of a mobile collaborative environment. It is only concerned with issues related to Human Computer Interaction (HCI). The collaborative aspects, essential in groupware systems, are not contemplated by this set.

There also two other GHE sets proposed in the literature: Baker's heuristics set (Baker *et al.*, 2001) and Bertini's heuristics set (Bertini *et al.*, 2011). However, neither covers all sides of a mobile groupware.

Baker *et al.* (2001) define a set of eight heuristics for evaluating the collaborative aspects in groupware systems. This set was based on conventional heuristics (i.e. Nielsen's heuristics). Although these heuristics are not specific to mobile devices, they present important issues that must be addressed in any groupware to facilitate collaborative aspects. Baker's heuristics cover three fundamental collaborative aspects: support for verbal and non-verbal communication, support for maintaining awareness, and support for group coordination.

Bertini *et al.* (2011) define eight heuristics specifically developed for evaluating mobile applications. This set of heuristics was also based on conventional heuristics evaluation methods, such as Nielsen's heuristics. This set of heuristics addresses the HCI and mobile aspects, but not in a collaborative perspective.

Some studies show Audience Response Systems developed to assist interaction of

participants in collocated groups. Among these systems, some have been developed for mobile devices (Lindquist *et al.*, 2007; Esponda, 2008; Scornavacca *et al.*, 2009; Teevan *et al.*, 2012). Others have been developed for web or desktop environments (Rechenthin and Molenda, 2009).

Rechenthin and Molenda (2009) developed an Audience Response Systems for using questionnaires in the classroom. Each student can answer a questionnaire individually and the teacher has access to the answers given by the students.

Teevan *et al.* (2012) developed an Audience Response System that allows the public to give feedback to the speaker. The system consists of three main components: mobile client who provides feedback; shared display to present the feedback; and widgets developed to include the speaker's feedback. The mobile client allows users to give positive or negative feedback for the lecture. The application is accessed via browser. The results collected by mobile clients are presented directly on a shared display. The widgets are designed to remind audience members to provide feedback and draw attention as interesting situations occur (large amount of positive or negative feedbacks, long period of inactivity, lots of participants, among others).

Esponda (2008) developed an Audience Response System to be used via mobile devices. The system allows that questions be spontaneously made at the time the questions arises (on the fly), but there is also the possibility of applying pre-designed questionnaires.

Scornavacca *et al.* (2009) present a system called TXT-2-LRN (text-to-learn), which uses SMS messages for students to answer the questions performed by the teacher during lessons. According to the authors, the use of the system increased the quality of student feedback and improved the interest and participation of students.

Lindquist *et al.* (2007) developed a system that allows students to submit exercise solutions in textual form or through pictures using their mobile devices.

Some authors develop models or frameworks in order to improve collaboration aspects in mobile applications. González and Ruggiero (2006) present a conceptual model of collaborative learning based on project execution. Silva and Rabelo (2012) present a decision support system based on a memorandum of decision and management support.

Audience response systems

In this section we present the two Audience Response Systems used to allow collaboration in classrooms, workshops, or conferences from mobile devices. These systems are used in collocated environments when a mediator interacts with the meeting participants. The mediator is the person responsible for organizing and monitoring the progress of the meeting. The presented systems are: Simple Question (Balestrin *et al.*, 2011) and Voting (Pereira *et al.*, 2013).

Simple Question

Simple Question has been developed to enable collaboration during preparation of questions in classroom lectures (Balestrin *et al.*, 2011). The questions can be asked through mobile or fixed hosts. The questions posed are available to other users, allowing participants, for example, to know in advance the questions to be asked to a speaker during a lecture.

The questions submitted are available to all participants. If a participant wishes to ask a question that is already registered, he/she has the option of supporting this question. Each question has a counter that is incremented at each participant support. And at the end, when the speaker ends his/her presentation, the tool shows all the questions in decreasing order of their number of votes.

In order to participate in the collaborative session it is necessary to use mobile devices with Wi-Fi or a computer connected to the network. The device must have a browser, through which the application could be used.

The user needs to be registered previously to use the application. Registration is useful to identify the question's authorship. And if the question is selected but not answered by lack of knowledge or lack of time during the session, it is possible for the speaker to contact the participant to answer the pending questions. However, in order to maintain the participants' identity secret during the session only the mediator is capable of identifying the questions' authors. Thus, it is supposed that secret authorship encourages the elaboration of questions by people who feel intimidated to ask a question in public.

As shown in Figure 1b, the questions are displayed as a list on the device. A title is assigned to each question by the participant, which is responsible for identifying it.

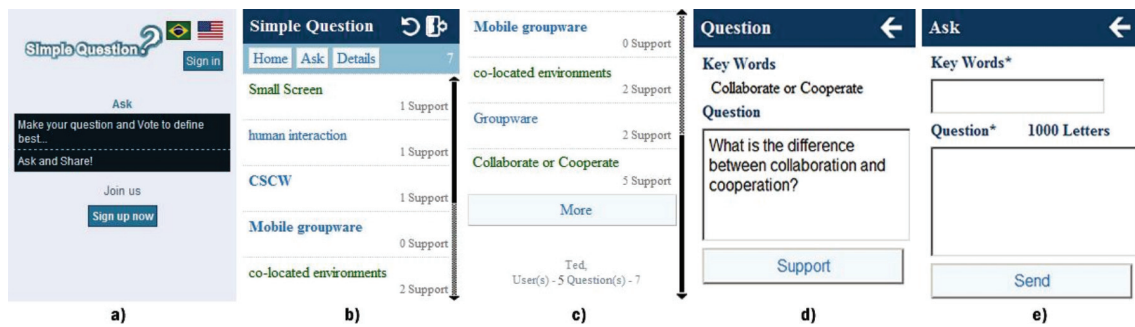


Figure 1. (a) Start screen; (b) Main screen without scrolling; (c) Main screen with scrolling; (d) “Question” screen; (e) “Ask” screen.

MEETING

Create


Edit

Open

Save

Close





Filter

Questions	7																																								
<table><thead><tr><th>Code</th><th>Title</th><th>Author</th><th>Votes</th><th></th></tr></thead><tbody><tr><td>1234</td><td>Collaborate or Cooperate</td><td>Martin</td><td>5</td><td></td></tr><tr><td>Specialist</td><td>Groupware</td><td>Olivia</td><td>2</td><td></td></tr><tr><td>Gabriel Balestrin</td><td>co-located environments</td><td>Karen Porter</td><td>2</td><td></td></tr><tr><td>Title</td><td>CSCW</td><td>Rosalie Chieli</td><td>1</td><td></td></tr><tr><td></td><td>human interaction</td><td>Brian Gliser</td><td>1</td><td></td></tr><tr><td>Date</td><td>Small Screen</td><td>Ted</td><td>1</td><td></td></tr><tr><td></td><td>Mobile groupware</td><td>Rosalie Chieli</td><td>0</td><td></td></tr></tbody></table>	Code	Title	Author	Votes		1234	Collaborate or Cooperate	Martin	5		Specialist	Groupware	Olivia	2		Gabriel Balestrin	co-located environments	Karen Porter	2		Title	CSCW	Rosalie Chieli	1			human interaction	Brian Gliser	1		Date	Small Screen	Ted	1			Mobile groupware	Rosalie Chieli	0		
Code	Title	Author	Votes																																						
1234	Collaborate or Cooperate	Martin	5																																						
Specialist	Groupware	Olivia	2																																						
Gabriel Balestrin	co-located environments	Karen Porter	2																																						
Title	CSCW	Rosalie Chieli	1																																						
	human interaction	Brian Gliser	1																																						
Date	Small Screen	Ted	1																																						
	Mobile groupware	Rosalie Chieli	0																																						
Duration	60 min.																																								
Registered Participants	6																																								

Figure 2. Mediator’s main screen.

The question posted last is shown at the top of the list. The list presents up to seven questions at the same time. This is because some mobile browsers are limited and cannot render a large amount of information in their windows. Thus, at the end of the list there is a button that allows the user to view the other questions, as shown in Figure 1c.

In order to access the question content it is necessary to click on its title. A new window displays the author’s name, the title of the question and its content. Once the question access is done, the user has the option of supporting it. Figure 1d shows the screen corresponding to this situation.

In the main screen, it is possible to see the number of questions available and the number of connected users. It has a button to access some information like the lecture presentation title, the speaker’s name, expected duration, among others features. Also, through this screen, the participant can: prepare questions,

as shown in Figure 1e; exit application, and refresh the screen data.

The mediator is the person responsible for organizing and monitoring the progress of the meeting. He/she can use a fixed computer, notebook, netbook or tablet to manage the application. Using the main screen shown in Figure 2, the mediator can create, edit, open, save and close sessions of questions. It is possible to filter the questions by title. This feature becomes more interesting when the speaker receives a large number of questions. All the questions submitted by participants are listed in descending order of support. So the questions that are on the top of the list represent those that obtained the largest number of supports.

The information about a question can be accessed by a click on the current line where it appears. Whenever a new question is shown in the user’s screen, it appears in bold. After viewing it, it loses the bold style. This helps the mediator differentiate the questions already

read from the unread ones. The mediator can also access a question and mark it as answered or unanswered. In case of answered questions, a green flag appears next to the question. The unanswered questions have a red flag.

As presented, the mediator has access to the information about the user who submitted the question. Thus, the mediator or the speaker may contact this user later.

Voting

The Voting tool has been developed to allow voting during lectures, meetings, conferences and classrooms, i.e., when it is necessary to vote in collocated environments (Pereira *et al.*, 2013). In collocated environments, the participants of collaboration are located physically close to each other. In these environments, a variety of communication forms are used to establish a shared understanding about the tasks performed during the collaboration. Communication may be performed by audio (speech, dialogue and voice intonation), and vision (gestures, facial expressions and positioning) or the environment (spatial relationships, presence and manipulation of objects). The Voting tool allows collaboration between mediator and audience when a question is made by the mediator. Instead of each participant raising his/her hand, he/she votes through their mobile devices. The system runs in a web browser, thus, it can be used in any device that has Internet access.

The questions to be voted are created by the mediator before or during lectures. These questions are available to the participants. They can vote in each question just once.

The votes are counted and all users can see the count of votes of each question. Furthermore, the responses generate a graphic that can be viewed by the users.

The tool also allows participants to create a new voting session. The participant can create a session to know something about the lecture or initiate a discussion about some subject. When the participant creates the session, this session should be accepted by the mediator. After the mediator accepts the voting session proposal all users can see the new session.

Figure 3 presents further details about the Voting tool. Figure 3a shows the system login. For the user's sign in in Voting it is necessary that an account be created. This is necessary to control the participants during the meeting, for example, if a participant causes any kind of confusion he/she can be banned from the meeting by the mediator. In the login screen, it is required to select the meeting among the list of activate meetings. When the user signs in, he/she will see all active votings of the meeting, as shown in Figure 3b. The user can vote or view the results, clicking on the voting title. When the user opens the voting, he/she will see a list of options to select one question and vote. Also, he/she can view the results clicking on the Results button as shown in Figure 3c. After voting or by clicking on the Results button, the user will see the partial results of voting in a graph as presented in Figure 3d.

When a participant votes, only his/her vote is counted, but the participant's identity is kept confidential. Thus privacy is ensured. Privacy is important for the participants not to feel constrained and to collaborate more.

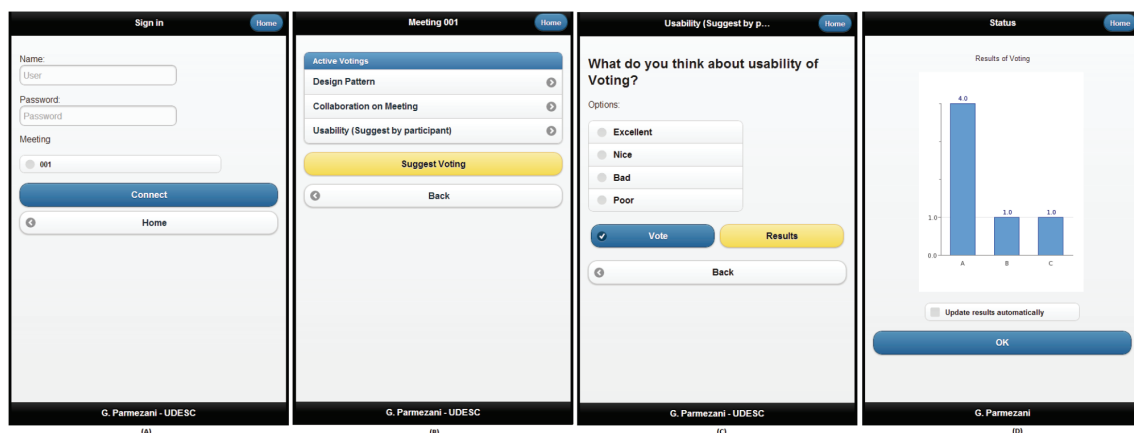


Figure 3. Participants' mobile interface: (a) Sign in the meeting; (b) Voting home page; (c) Participants' voting; (d) Results of voting.

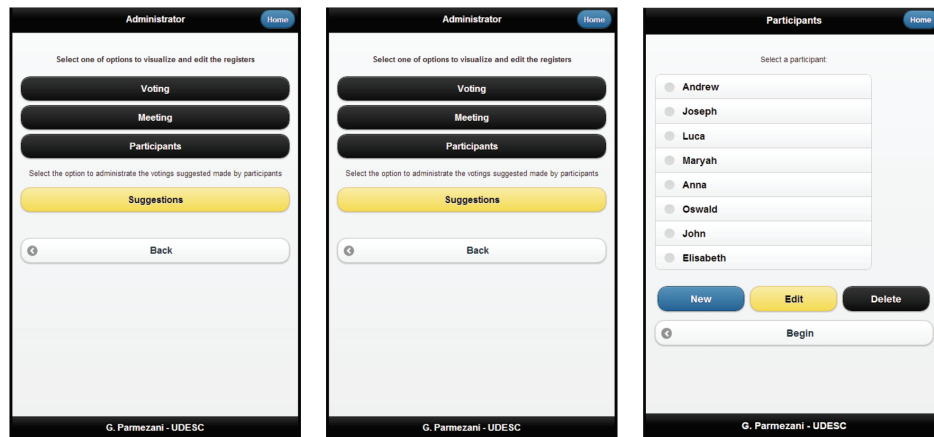


Figure 4. Mediator's mobile interface.

Figure 4 presents the mediator's screen. Mediator is the person who controls the voting, the meeting, the participants, and the voting suggestions made by participants. To create or delete a voting, the mediator should click on the Voting button. To enable and disable meetings, the mediator should click on the Meeting button. And to control the participants, the mediator should click on the Participants button. Thus, the mediator can control all the meeting and provide collaboration during the lecture.

Heuristics for mobile groupware evaluation

A heuristic evaluation is a usability inspection method in which the evaluation is based on a set of guidelines and desirable characteristics that describe the interaction, guiding the evaluators to systematically inspect the interface for problems that affect usability (Nielsen and Mack, 1994).

Conducting a heuristic evaluation consists in analyzing the interface to report problems, according to the heuristics and the expertise of the evaluators, seeking inconsistencies in the interface about the principles of usability. In this method experts examine the system and diagnose problems and barriers that users are likely to encounter during the interaction (Barbosa and Silva, 2010). Between three and five evaluators for evaluation are recommended.

Baker *et al.* (2001) report that the groupware evaluation is more difficult than conventional systems, because conventional usability evaluation methodologies do not allow discovering specific problems of Computer Supported Cooperative Work (CSCW) environments.

In order to achieve the main goal of this paper we employed a set of heuristics that include points to evaluate mobile groupwares. The heuristics used are based on the following sets: Nielsen's, Baker's and Bertini's heuristics. The three sets were chosen because they evaluate the usability in mobile groupwares and collaboration in groupwares. In order to define the heuristics to be used in this work, we joined all the three sets and we removed the duplicate heuristics.

As shown in Figure 5, the three heuristics sets are complementary to evaluate the mobile groupware usability: Nielsen's heuristics (Nielsen and Mack, 1994) evaluate the usability of conventional systems (HCI aspects); Baker's heuristics (Baker *et al.*, 2001) evaluate issues related to collaborative aspects, and Bertini's heuristics (Bertini *et al.*, 2011) evaluate issues related to mobile devices.

After studying the three sets we found the following 15 heuristics:

H1 - Visibility of system status. The system should provide appropriate feedback to users within reasonable time (Nielsen and Mack, 1994). Through mobile devices, the system should keep the user informed about what is happening. Moreover, the system should prioritize messages about critical and contextual information, such as network status, battery and environmental conditions (Bertini *et al.*, 2011). The system should provide awareness information about what is going on, who is in the workspace, where they are, and what they are doing (Baker *et al.*, 2001).

H2 - System capability with the real world. The system should use terms that are familiar to the user rather than software-ori-

can do before requiring discussion, instruction or consultation with others (Baker *et al.*, 2001).

H13 - Allow people to coordinate their actions. One of the problems found during face-to-face situations is how group members mediate their interactions. Coordinating actions involves performing the tasks in the right order, at the right time, and without bypassing the restrictions imposed (Baker *et al.*, 2001).

H14 - Provide appropriate methods for group communication. In collaborative environments people communicate, and thus, build a common understanding about the collaborative tasks (e.g. exchange ideas, discuss, negotiate, align ideas, and make decisions) (Gerosa *et al.*, 2003). The groupware should provide two media (Baker *et al.*, 2001): verbal communication (e.g. chat, video, voice); and nonverbal communication (e.g. awareness of the others' actions over the shared artefacts).

H15 - Facilitated group management. One problem found in groupwares is how the participants begin the collaborative section. The system should present information to facilitate the group management (Baker *et al.*, 2001): being available; aware of others participants, who is available for interaction; contacting others; and working together.

Evaluation methodology

Each problem found according to the heuristics should be associated with a severity degree that is based on a combination of three factors (Kimura *et al.*, 2012): (i) the frequency with which it occurs (e.g. it is common or rare); (ii) the impact of the problem when it occurs (e.g. it is easy or difficult to overcome it); (iii) the persistence of the problem (e.g. it is a problem that occurs only once and the user can overcome it as long as he/she knows it exists, or if the users are repeatedly bothered by it).

These factors influence the severity levels used in the evaluation, which can be classified as (Nielsen, Mack, 1994): (0) it is not necessarily seen as a usability problem; (1) cosmetic problem that does not need to be fixed, unless there are resources and time available; (2) minor usability problem, with low priority to be fixed; (3) major usability problem, with high priority to be fixed; (4) catastrophic usability problem, where it is imperative to fix it.

The evaluation was performed using two Audience Response System, Simple Question and Voting, as presented in "Audience Response System". The evaluation was con-

ducted by three evaluators, all with prior knowledge of the evaluation method and environments to be evaluated. The evaluation focused on functionalities and on collaboration. Then, the difference between devices does not affect the evaluation because they allowed to evaluate these items. The Simple Question evaluation was conducted only over the mobile participants interface (the mediator interface was disregarded from the study), accessed via browser on an Android operational system. One mobile device Samsung Galaxy Tab P1000L and two Android 2.3 emulators were used for the assessment. The Voting evaluation was conducted similarly as in Simple Question evaluation. The study was conducted only over the mobile participants interface, accessed via browser on an Android operational system. Two mobile devices (one LG Nexus 4 with Android 4.4 and one Samsung Galaxy GT I5500B with Android 2.2) and one Android 2.3 emulator were used for the assessment.

The evaluation procedure was performed in three steps, as shown in Kimura *et al.* (2012): (i) initial exploration of the system, seeking and understanding the features to be evaluated; (ii) evaluation period, when each evaluator used the Simple Question and Voting systems separately, inspecting their interface; (iii) consolidation of the evaluation, where evaluators identified all the problems found, discussing their severity and suggesting solutions.

After the evaluation, the following pieces of information were identified: (i) usability problems found; (ii) heuristics violated; (iii) severity associated with each problem; (iv) possible solutions suggested.

Evaluation results

The results of the heuristic evaluation for each of the evaluated tools (Simple Question and Voting) are described below. Figure 6 presents the amount of problems found by the evaluator. The heuristic evaluation found 13 usability problems. Two are common problems with the Voting tool.

Simple Question

The heuristic evaluation found 13 usability problems. Two are common problems with the Voting tool. Each one of the evaluators found a different set of problems, as described in Figure 3a. The distribution of the problems found, as shown in Table 1, indicates that the

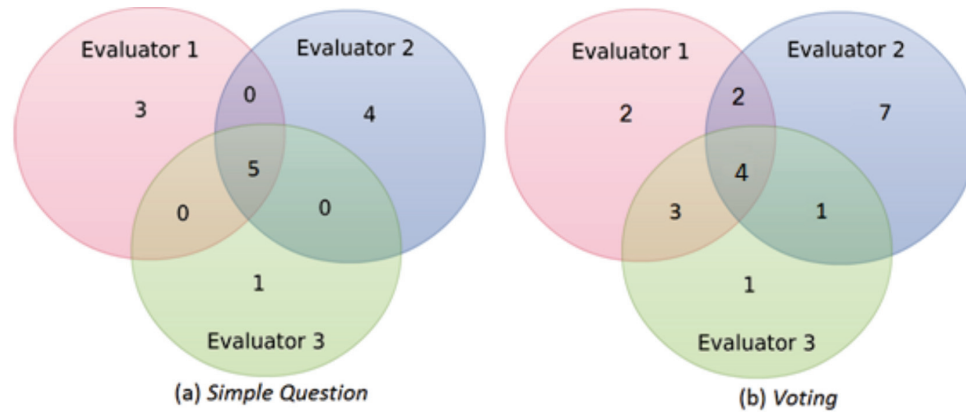


Figure 6. Amount of problems found by the evaluator.

main problems of Simple Question are related to aspects of Error prevention (H7) and Ease of input, viewing, and screen reading (H8). No problem regarding cosmetic problems, with low gravity of correction (severity level 1) or considered a catastrophic usability problem (severity level 4) were found. Another aspect that can be observed is that for heuristics H4, H9, H10 and H11 we did not find usability problems, which shows that the Simple Question application conforms to the one advocated by these heuristics.

The main problems found in Simple Question were:

Problem #1: Error messages are displayed on the user's interface without any treatment form (e.g. SQL errors). This problem violates H7 and it is classified as severity level 3. As a solution, it is suggested to handle the SQL errors. If it is necessary to inform the user about the errors occurrence, it should be presented in a natural language to the user.

Problem #2: The maximum size of a question on the Simple Question is 1000 characters. However, when users ask a longer question than allowed, no error is reported to the user. Furthermore, a message is displayed informing that the question was registered successfully, but the question is not listed as a question asked. This problem also violates H7, and it is classified as severity level 3. As a solution, it is suggested doing the treatment of the field limit, warning the user about the established limit and allowing him/her to do the necessary adjustments in the question.

Problem #3: Lack of feedback to the user about the system status and about the actions performed by him/her (e.g. what questions did he ask? What questions did he support?

What questions did he evaluate?). At the design phase, we decided that the supported questions should be indicated in green in the screen. This problem violates H1 and H3, and it is classified as severity level 3. As a solution, it is recommended to display a legend indicating the meaning of colors and icons, to make the system status clear to the user.

Problem #4: Haziness in identifying the question assessed, because the same icon is displayed on the interface both if the user likes or dislikes the question (e.g. the icon [LIKE] is displayed to the questions asked, unanswered and not understood). This problem violates H1 and H3, and it is classified as severity level 3. As a solution, it is suggested using different colors or icons to questions understood and not understood (e.g. positive icon/green color - to questions understood; negative icon/red color - to questions not understood).

Problem #5: The system allows the user to initiate a blocked action, making the process logic confusing (e.g. by clicking on a question already supported, the support button is not locked. The system informs the user that the question has already been supported by him only after he tries to support the question again). This problem violates H2 and it is classified as severity level 2. As a solution, in addition to the colors legend, the support button should be locked (since the action can no longer be performed by him).

Problem #6: When the user alternates between the system screens the previous configuration is lost (e.g. order of the questions). This problem violates H5 and it is classified as severity level 2. As a solution, it is advised to keep the ordering chosen by the user when he/she comes back to the initial context.

Problem #7: Inadequate handling error and loss of user data (e.g. when the user uses a title which already exists in a new question, the system handles the error but loses all information previously entered in the form). This problem violates H7 and H8 and it is classified as severity level 3. As a solution, it is suggested keeping all information already entered, only clearing the title field so that it is entered again.

Problem #8: Lack of questions pagination (e.g. options to return to the previous page, go to the next page). This problem violates H8 and it is classified as severity level 3. As a solution, it is suggested allowing question pagination and allowing the navigation through the question list.

Problem #9: The information update must be performed manually (e.g. questions list, participants, supports number). If the users do not constantly update the application, their interaction will be impaired because they will not have access to the new questions. This problem violates H8 and it is classified as severity level 2. As a solution, it is suggested running the data update automatically, to maintain the environment integrity visible to the user.

Problem #10: Lack of search engine queries (for example, the system does not allow to search by title or keyword on questions already completed). In this way, the user does not see if a particular question has been asked, hampering the collaboration process. This problem violates H12 and it is classified as severity level 3. As a solution, it is suggested allowing finding questions by title or keywords.

Problem #11: There is no separation between questions already answered and unanswered questions (e.g. the questions are ordered by default configuration, and there is not any distinction for the questions). This problem violates H6 and it is classified as severity level 2. As a solution, the definition of separated places for answered and unanswered questions is suggested, improving the aesthetic design of the application.

Voting

The heuristic evaluation found 20 usability problems, two are common problems with the Simple Question tool. Each of the evaluators found a different set of problems, as described in Figure 3b. The main problems of Voting are related to the aspects of Consistency and Standards (H3), Visibility of system status (H1), Sys-

tem capability with the real world (H2) and Ease of input, viewing, and screen reading (H8). Problems found in Voting include both low-gravity correction problems (severity level 1) and catastrophic usability problems (severity level 4). In general Voting has presented more problems than Simple Question. The worst evaluated heuristic was H3, with six problems, and H2, with two catastrophic problems. Only one heuristic (H6) showed no problems. For collaborative aspects (H10 to H15) all of them have at least one problem, which indicated the need of Voting for improvements in the aspects of collaboration.

The main problems found in Voting were:

Problem #1: The system does not update the questions suggested by the participants automatically. It is necessary for the participant to refresh the screen to view the new questions. This problem violates H1 and it is classified as severity level 3. As a solution, it is suggested updating automatically the questions in a reasonable time so that the user does not become outdated.

Problem #2: The system lets voting again for the question already voted. The user is not notified that he has already voted. This happens just after he votes again. This problem violates H4 and H7, and it is classified as severity level 2. As a solution, it is recommended that the question already voted be marked or that the vote button be blocked if the user has already voted for the question.

Problem #3: When the participant suggests a question, it is not clear what happened with his/her suggested question and whether it was accepted or not. The user does not know what happened. The problem violates H1 and it is classified as severity level 2. As a solution, an appropriate feedback to user is suggested, stating that the question is being evaluated by an administrator, so it will be clear to the user what happened to the question suggested by him.

Problem #4: The system does not allow changing the profile or different settings by users. The problem violates H5 and it is classified as severity level 0. As a solution, it is suggested allowing the user configuration flexibility to facilitate the system use (e.g. differentiating it for experienced and novice users).

Problem #5: When suggesting a question, the system does not define the minimum and the maximum size of characters in the fields, nor does it validate the question, allowing the user to enter any characters in the fields or to leave them blank. This problem violates H3

and H7, and it is classified as severity level 3. As a solution, it is suggested validating the size of the fields, avoiding mistakes and not allowing questions with blank data (e.g. title, options).

Problem #6: The system lets the user click several times to submit and submits N times the same question to the server. This problem violates H3 and H7 and it is classified as severity level 2. Providing proper feedback to the user that the question has been sent for validation of an administrator avoids N clicks to submit the same question.

Problem #7: Security issue with the password because when creating the account the password is displayed to the user. The problem violates H9 and it is classified as severity level 3. As a solution, it is recommended encrypting the password and sending an email to inform its recovery if the user forgets it.

Problem #8: It does not show who created each question (anonymous) to maintain privacy. However, it is shown when others have suggested questions and voted. This problem violates H10 and it is classified as severity level 1. As a solution, it is suggested providing information on who created the question in addition to other information available.

Problem #9: By suggesting questions, may conflict with the answers suggested by the participants (e.g. a question from a participant with the answer options suggested by another participant). The problem violates H11 and it is classified as severity level 3. As a solution, it is suggested providing mechanisms to avoid conflicts between the participants' actions.

Problem #10: There is no functionality for communication among participants through the system due to a co-located environment (lecture). However, it is a collaborative system then it should provide mechanisms for allowing communication among the participants. This problem violates H12 and it is classified as severity level 2. As a solution, it is recommended allowing communication between the participants in the system.

Problem #11: The DEL button is not enabled to delete unwanted options when creating a question. The problem violates H3 and H8 and it is classified as severity level 3. As a solution, it is suggested enabling the DEL button.

Problem #12: The BACK button does log off on the screen of suggested questions and on the screen with list of questions (navigation). The problem violates H2 and it is classified as severity level 4. This error is quite

serious because the user is disconnected from the system, and is unable to continue his/her activities. As a solution, it is suggested that the BACK button return to the previous screen, according to the logic of the process.

Problem #13: The HOME button does log off and shows no warning/question asking the user for confirmation. The problem violates H2 and H3, and it is classified as severity level 4. This error, as error #12, is very serious because the user is disconnected from the system, and is unable to continue his/her activities. As a solution, it is suggested that the HOME button return to the main screen of the application (list of questions).

Problem #14: Lack of sorting and filtering options in the list of questions screen. The problem violates H8 and H9 and it is classified as severity level 2. As a solution, it is suggested that the navigation screen of the questions have sorting and filtering options of questions to facilitate entry and viewing of the data, and promoting the improvement in the aesthetic design of the application.

Problem #15: It does not show information about how many people had voted for each question. The problem violates H1 and H10 and it is classified as severity level 1. As a solution, it is suggested that each question present the number of people who voted on it, thus providing an updated feedback from system status to the user and allowing the user to view the activities of other participants.

Problem #16: Inconsistency of different names for the system buttons (BACK, END and HOME) that go to the same place. The problem violates H3 and it is classified as severity level 3. As a solution we encourage BACK, HOME and END buttons to perform their respective task, or having the same name if the goal is to perform the same task. For example, if BACK and END log off the system, so they should be called LOG OUT or EXIT.

Problem #17: On the results screen, the charts show data without proper subtitles (e.g. columns A, B, C). The problem violates H4 and H8 and it is classified as severity level 2. As a solution, it is suggested that the results graph show in the subtitles its options available to answer in every question, facilitating the recognition and visualization of data to the user.

Problem #18: Inconsistent initial navigation. It should check the LOGIN user type. The system first asks the user profile and then asks the user/password, while could inform the user, password and profile. The system

also recognizes the profile entered by the user (through the registration done previously). The problem violates H2 and H3 and it is classified as severity level 1. As a solution, it is suggested that the system recognize and automatically bring the user profile through the registration done previously.

Common problems

The heuristic evaluation found two common usability problems in both tools. These problems were identified by the three evaluators and they are described in Table 1 with letters A and B. The common problems are related to coordination (H13), group communication (H14), and group management (H15), as follow:

Problem #A: The system does not provide appropriate methods for group communication. The system only provides non-verbal communication media between participants (e.g. provides awareness about the actions of others over the shared artefacts -- in the Simple Question case, only awareness information about the questions status is provided; in the Voting case, only awareness information about the voting when the user selects a vote in the list is provided). However, as Baker *et al.* (2001) comment, in a groupware it is also necessary to provide means of verbal communication (e.g. features to discuss or leave comments). This problem violates H14 and is classified as severity level 2. As a solution, new features to enable verbal communication, debate and which allow leaving comments are suggested.

Problem #B: The system does not allow group management (coordination). Neither tool provides appropriate ways for the participants to initiate a collaborative section and then to coordinate the group activities during the tool use. The collaborative section is initiated only by a mediator, and only he/she has the control over the shared artefacts in both cases. The participants have to wait for the mediator to accept/reject a question/voting proposed, so they can keep the collaboration (e.g. voting). Furthermore, the user experience in the collaborative system is inherently associated with the mediator abilities. If the mediator plays his/her role well (e.g. he/she does not forget to accept/reject the submitted questions/voting) participants will have good chances to collaborate effectively, otherwise (e.g. mediator cannot efficiently coordinate or does not answer

in a short time all requests) the group will have difficulty using the tool and improving collaboration. This problem violates H13 and H15, and it is classified as severity level 2. As a solution, it is suggested that the system make mechanisms available to facilitate the coordination and management of the group (e.g. list of active participants, presence information) and mechanisms of communication.

Final remarks and conclusion

This work presents a set of 15 heuristics established to evaluate the usability for mobile collaborative applications. These heuristics were based on three sets of heuristics presented in the literature: Nielsen's heuristics set (Nielsen and Mack, 1994) to assess the usability of conventional applications, especially to assess the Human Computer Interaction (HCI) aspects; Baker's heuristics set (Baker *et al.*, 2001), developed to assess collaborative aspects; and Bertini's heuristics set (Bertini *et al.*, 2011), developed to assess the usability of mobile applications. With the set of heuristics raised in this study, it was possible to evaluate the usability and efficiency of two mobile collaborative systems: Simple Question and Voting.

For Simple Question, among the 15 heuristics established, some have not found usability problems: recognition rather than recall (H4); aesthetic, social and private conventions (H9); providing communication of shared artifacts (H10); and providing protection (H11). This indicates that Simple Question is in accordance with what is expressed by these heuristics. As shown in the heuristic evaluation, 6 of 13 issues raised are related to error prevention (H7), and ease of input, viewing, and screen reading (H8). It is indicated that the application needs to improve its interface with regard to HCI aspects (Nielsen heuristic based) and mobile devices utilization (Bertini heuristic based).

For Voting, just H6 has not found problems and 12 of the 20 problems are related to the aspects of: consistency and standards (H3); visibility of system status (H1); system capability with the real world (H2); and ease of input, viewing, and screen reading (H8). Evaluation found some problems with severity level 4, which indicates the system needs improvements in these aspects. Furthermore, both systems had problems regarding the heuristics: allowing people to coordinate their actions

(H13); providing appropriate methods for group communication (H14); and facilitated group management (H15). These heuristics in particular indicated the need for improvement in aspects of communication and coordination to fully enable the collaborative aspects.

Based on the results and experiences obtained over the heuristic evaluation performed, it is suggested that the heuristics listed in this paper is applicable to evaluate mobile groupware usability, addressing HCI, mobile and collaborative aspects. The requirements analysed in Simple Question and Voting indicated that both are in accordance with some usability heuristics. However, some common problems have also identified such as inappropriate communication methods, low coordination and group management. This shows that there are details to be improved in both systems. Although in both cases the systems are used in collocated environments, there is still a need to provide adequate media for communication, collaboration, and group management. The employed set of heuristics can also be used during the development of mobile groupware for evaluating it during its construction.

Baker et al. (2001) have adapted Nielsen's heuristic evaluation method to groupware systems. Although they believe that their heuristics could be applied to groupware systems, they did not develop an evaluation in a real scenario. In this work, we applied the heuristics proposed by Baker in different scenarios. We found that these heuristics helped us to focus our attention on the critical issues related to the analysed groupwares. The evaluators did not have problems in finding a reasonable number of usability problems. In addition, we believe that this work can help to prove that the proposed heuristics are useful for evaluating groupwares. Similarly to Bertini et al. (2011), we also performed a set of heuristics for two mobile applications, however, Bertini et al. (2011) have used 8 usability experts, while we used a smaller number of evaluators (3). We also believe that the experimental study conducted in this work is important for consolidating Bertini's heuristics for the evaluation of mobile applications.

As future work we intent to carry out more evaluations in other groupwares to find the consolidations of the employed set. We also aim to evaluate groupwares with other consolidated methods and then compare the results with the employed set of heuristics.

References

- ANTUNES, P.; HERSKOVIC, V.; OCHOA, S.F.; PINO, J.A. 2012. Structuring dimensions for collaborative systems evaluation. *ACM Computing Surveys (CSUR)*, **44**(2):8.
<http://dx.doi.org/10.1145/2089125.2089128>.
- BAKER, K.; GREENBERG, S.; GUTWIN, C. 2001. Heuristic Evaluation of Groupware Based on the Mechanics of Collaboration. In: Conference on Engineering for Human-Computer Interaction, 8th, Calgary, 2001. *Proceedings...* Calgary, **2254**:123-139.
http://dx.doi.org/10.1007/3-540-45348-2_14.
- BALESTRIN, G.; BERKENBROCK, C.D.M.; HIRATA, C.M. 2011. Uma Ferramenta de Colaboração Móvel para Auxiliar a Interação em Palestras Presenciais. In: Simpósio Brasileiro de Sistemas Colaborativos, 8th, Rio de Janeiro, 2011. *Proceedings...* SBSC, p. 76-82.
- BARBOSA, S.D.J.; SILVA, B.S. 2010. *Interação Humano-Computador*. 1^a ed., Rio de Janeiro, Elsevier, 408 p.
- BERTINI, E.; CATARCI, T.; DIX, A.; GABRIELLI, S.; KIMANI, S.; SANTUCCI, G. 2011. Appropriating Heuristic Evaluation for Mobile Computing. In: J. LUMSDEN (ed.), *Human-Computer Interaction and Innovation in Handheld, Mobile and Wearable Technologies*. Hershey, PA: Information Science Reference, p. 20-41.
- ESPONDA, M. 2008. Electronic voting on-the-fly with mobile devices. *SIGCSE Bull*, **40**(3):93-97.
<http://dx.doi.org/10.1145/1597849.1384298>.
- GEROSA, M.; FUKS, H.; LUCENA, C. 2003. Suporte à percepção em ambientes digitais de aprendizagem. *Revista Brasileira de Informática na Educação*, **11**(2):75-85.
- GONZÁLEZ, L.A.G.; RUGGERIO, W.V. 2006. Modelo Aprendiz para atividades colaborativas de projeto em Sistemas de Aprendizagem Eletrônico. *IEEE Latin America Transactions*, **4**(4):285-290.
- HERSKOVIC, V.; PINO, J.A.; OCHOA, S.F.; ANTUNES, P. 2007. Evaluation methods for groupware systems. *Groupware: Design, Implementation, and Use*, p. 328-336.
http://dx.doi.org/10.1007/978-3-540-74812-0_26.
- KIMURA, M.H.; MANTAU, M.J.; KEMCZINSKI, A.; GASPARINI, I.; BERKENBROCK, C.D.M. 2012. Usability evaluation of Facebook's privacy features: comparison of experts and users. In: IADIS International Conference WWW/Internet, Madrid, 2012. *Proceedings...* Madrid, **1**:311-318.
- LINDQUIST, D.; DENNING, T.; KELLY, M.; MALANI, R.; GRISWOLD, W.G.; SIMON, B. 2007. Exploring the potential of mobile phones for active learning in the classroom. *SIGCSE Bull*, **39**(1): 384-388.
<http://dx.doi.org/10.1145/1227504.1227445>.

- NIELSEN, J.; MACK, R.L. 1994. *Usability inspection methods*. 1^a ed., United State of America, John Wiley & Sons, 448 p.
- PEREIRA, G.P.; BERKENBROCK, C.D.M.; BERKENBROCK, G.; HIRATA, C.; ARAÚJO, L.P. 2013. Uma ferramenta para apoiar votações em ambientes colocalizados. In: Seminário de Computação, Santa Catarina, 2013. *Anais...* Santa Catarina, **22**:9.
- PINELLE, D.; GUTWIN, C. 2000. A review of groupware evaluations. In: International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, 9th, Saskatoon, 2000. *Proceedings...* IEEE, p. 86-91.
<http://dx.doi.org/10.1109/ENABL.2000.883709>.
- RECHENTHIN, M.; MOLEND, P. 2009. *Student Response Systems*. Loyola University Chicago, Chicago, 5 p.
- SCORNAVACCA, E.; HUFF, S.; MARSHALL, S. 2009. Mobile phones in the classroom: if you can't beat them, join them. *Communications of the ACM*, **52**(4):142-146.
<http://dx.doi.org/10.1145/1498765.1498803>.
- TEEVAN, J.; LIEBLING, D.; PARADISO, A.; SUAREZ, C.G.S, VON VEH, C.; GEHRING, D. 2012. Displaying mobile feedback during a presentation. In: International conference on Human-computer interaction with mobile devices and services, 14th, California, 2012. *Proceedings...* MobileHCI, p. 379-382.
<http://dx.doi.org/10.1145/2371574.2371633>.
- SILVA, M.V.D.; RABELLO, R.J. 2012. A semi-automated distributed decision support system virtual enterprises. *IEEE Latin America Transactions*, **10**(1):1235-1242.
<http://dx.doi.org/10.1109/TLA.2012.6142467>.

Submitted on January 3, 2014

Accepted on June 12, 2014