SINGRAR Usability Study

Isabel L. Nunes[1,2] and Mário Simões-Marques[3]

[1] Faculdade de Ciências e Tecnologia - UNL, 2829-516 Caparica, Portugal
[2] UNIDEMI, Faculdade de Ciências e Tecnologia - UNL, 2829-516 Caparica, Portugal
[3] CINAV, Portuguese Navy, BNL/Alfeite, 2810-001 Almada, Portugal
imn@fct.unl.pt, mj.simoes.marques@gmail.com

Abstract. Usability is a very important issue that affects the effectiveness and success of systems. Such importance becomes particularly critical when systems are complex, and when the accuracy and timeliness of operation is decisive to the system outputs. Naturally, the usability of decision support systems used for emergency management is of utmost relevance. The present paper addresses a usability study performed to the Portuguese Navy SINGRAR system.

Keywords: emergency management system, usability study, SINGRAR.

1 Introduction

SINGRAR is the acronym of the Priority Management and Resource Assignment Integrated System. This system is a distributed expert system used on board of Portuguese warships, to compile information about the status of the platform, equipment and personnel, supporting the management of emergency situations (e.g., fires, floods) and giving advice about recommended courses of action (e.g., damage control, equipment repair priorities). SINGRAR is an emergency management system (EMS) used in combat and emergency, situations in which users may have to wear personnel protective equipment, e.g. gloves, anti-flash gear or masks. This equipment presents a challenge in the operation of the system [1].

The importance of usability becomes particularly critical when systems are complex, and the accuracy and timeliness of operation is decisive to the system usefulness. Naturally, the usability of EMS is of utmost relevance.

Usability is a critical aspect to consider in the development cycle of software applications, and for this purpose, user-centered design and usability testing must be conducted. The design and testing cannot ignore the context of use of software, whose knowledge is essential. Usability of a system is characterized by its intuitiveness, efficiency, effectiveness, memorization and satisfaction. Good usability allows decreasing the time to perform tasks, reducing errors, reducing learning time and improving system users' satisfaction [2].

Usability assessment can be done based on objective and subjective evaluations. Objective evaluation of performance, measures the ability of users to operate the system. The subjective assessment of users' preferences evaluates how much users like the system. ISO 9241 - Part 11 refers that usability is measured as a function of the degree to which the goals of the system are achieved (effectiveness), of the resources

A. Marcus (Ed.): DUXU/HCII 2013, Part I, LNCS 8012, pp. 359–368, 2013.

(such as time, money, or mental stress) that must be spent to achieve the objectives (efficiency) and of the extent to which users of the system find it acceptable (satisfaction) [3]. Therefore, effectiveness, efficiency and satisfaction are criteria that have to be considered when evaluating the usability of a system. To assess these criteria, it is necessary to consider sub-criteria, which are measurable.

There is a wide range of tools and methodologies for identifying and evaluating the usability of a system, thus contributing directly or indirectly, for its improvement. The selection of these tools and methodologies depends on the objective to achieve, which usually is related to the development phase the system is in. Some approaches are better suited to the design stage (e.g., analysis of context of use and task analysis), while others are more suited to early stages of development and prototyping (e.g., brainstorming, prototyping) and others for the evaluation and testing (e.g., analytical and heuristic evaluations, SUMI). A compilation of methods can be found in [4].

For instance, SUMI is a rigorously tested and validated questionnaire based method to measure software quality from a user's perspective [5]. This tool is supported by an extensive reference database and embedded in an effective analysis and reporting tool, has been applied to a great number of projects.

The goal of the study was to identify the usability factors affecting operators' performance, recommend potential solutions to improve SINGRAR, and assess the gains achieved by the implementation of improvements introduced during the study period.

2 SINGRAR Characterization

SINGRAR is used to generate a common operational picture of the ship and their systems; support the decision making regarding courses of action and human resources allocation in critical combat and emergency situations; and support the flow of information between the different technical and operational areas of the ship. The compilation of a consolidated picture is made by operators from different technical areas, and information is automatically shared among the EMS terminals connected to a local area network.

Figure 1 illustrates examples of interfaces that combine graphical, tabular and text information used to characterize the environment, the status of the ship and their equipment and to support the decision-making process.

Despite some complexity of the interfaces, the operation is relatively intuitive for users who are familiar with the tasks that the system supports. The system was designed to maximize compatibility with existing procedures and the previously existing manual recording media, to allow a simple transition from one method to the other.

Users - SINGRAR is used by military personnel onboard Portuguese Navy ships. Users can be categorized into two groups, decision makers (Captain and Heads of Department) and technicians.

Tasks - Tasks performed with SINGRAR depend on the profile of users, which are related to their responsibilities within the organization on board. Usually the operating requirements are inversely proportional to the user level within the chain of command. That is, the activity performed by the Captain is low, increases slightly for the other decision makers, and is maximal for the technicians.

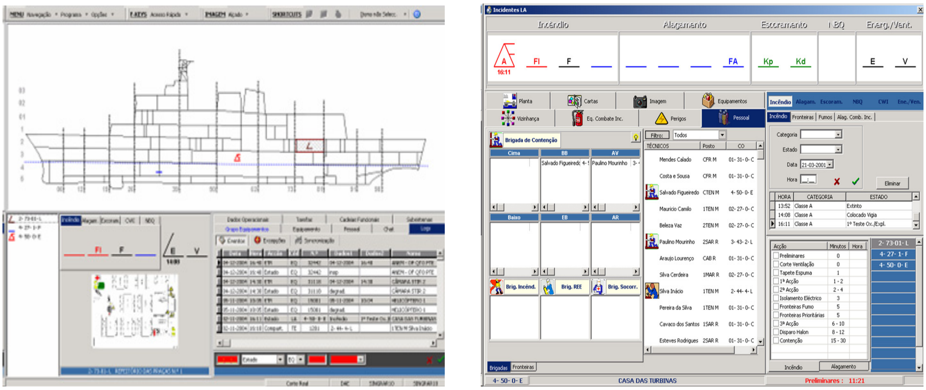


Fig. 1. SINGRAR - examples of interfaces

Operational environment - SINGRAR can be used in all degrees of readiness of the ship, from the ship moored alongside to Battle Stations. The number of workstations required increases as the ship readiness evolves towards Emergency or Battle Stations, which is the highest.

SINGRAR is installed mainly in desktop computers, but runs also in portable computers. Portable computers are used, for example, in Damage Control command posts. In these posts operators are standing and working in a confined area that usually does not allow the use of a mouse or a trackball. These workstations are not suitable for an extended operation of the system. Desktop based workstations are more suitable since they allow a more comfortable use of the computer, a factor which is very important to ensure an efficient, effective and satisfactory use of any type of application.

3 SINGRAR - Usability Evaluation Methodology

The usability study was prompted after users of the EMS advanced prototype reported some difficulties in inputting data at an adequate pace. These difficulties were experienced especially when dealing with high tempo incidents while training demanding operational and emergency scenarios.

The EMS was developed from the very beginning with usability concerns, particularly because the system is operated using a quite high number of interfaces, some of them presenting a high density of information which is complex in nature. In fact, the use of critical applications, developed based on information technologies, such as the EMS, had an exponential increase in recent years. However, their development is challenging considering that these systems are required to have high reliability and their users have to be able to use them effortlessly, with minimal training and that the tasks have to be performed in the shortest time possible. Thus, it is extremely important to ensure that usability is taken into consideration since the early stages of the development of such systems.

The methodology of the usability study considered three phases: data collection, results analysis and recommendations.

During data collection a group of users was observed directly by the analyst while operating the EMS performing activities defined in test script that reproduces typical operation situations. The work sessions were performed by crew members using the system both in a simulator ashore and onboard ships. The data collection procedure was designed to involve a significant sample of users within the target population, performing tasks associated with the operation of the system. The users that participated in the study were heterogeneous, either in terms of operation experience and technical expertise. To verify if the peculiarities of the work environment could affect the reliability of the usability study, some sessions were performed with the operators wearing the personnel protective equipment used in emergency, namely the gloves.

Data collection took place in 12 sessions, where users had to operate the system performing a set of tasks listed on a predefined and validated script. The script included 9 activities composed of 10 tasks each. The data collection procedure was designed to evaluate the efficiency, effectiveness and satisfaction of system operation and also to compare different operational methods (Table 1). The first two characteristics were assessed using objective data collected during usability tests. In addition, users answered the Portuguese version of SUMI [6], providing measures regarding the intuitiveness and ease of memorization of the application.

To collect and process the subjective data it was used the SUMI method, which employs a metric to assess the overall satisfaction or overall usability of software. The evaluation of the software quality from the users' point of view is based on a questionnaire with 50 statements, which was answered by the users at the end of each session.

Table 1. Characteristics assessed in the SINGRAR usability study

	SUMI Questionnnaire (Subjective Data)	Activity Analysis	
		Objective Data	Parameters
Effectiveness	X	X	Number of user errors Number of tasks finished in a given period of time
Efficiency	X	X	Number of actions performed Average, maximum and minimum time for performing the activity tasks
Satisfaction	X		-
Intuitiveness	X		-
Ease of memorization	X		-

For the analysis of objective data it were considered, first, the recording of the data inserted by each user and, second, the video recordings of the session in order to understand the circumstances in which the session evolved, and the reason for any disparate performance (e.g., long execution times and errors). The observation of the video recordings was particularly useful in isolating aspects of the procedures adopted to pass information to the users and of the methods of operation that proved to be problematic or, on the other hand, which constitute good practices to follow.

The records of the Event Logs exported from the EMS were processed to extract the main data elements (e.g., duration of each task), to detect errors in the input of the information defined by the script, and also to compare the progression of events with the desired state.

The video recording involved two video cameras and was intended to support the analysis of data collected, documenting the actions, comments and attitudes of users towards the application. The first recorder was placed in a fixed position perpendicular to the operator, and recorded actions, facial expressions and body posture of the users. The second camera was mobile collecting images of the computer screen, and recorded the actions performed by the users during the procedure.

Individual data were later aggregated in order to have a perception on the use of the system considering a broader set of users. In general, the data aggregation was based on average, minimum and maximum functions, which can identify trends and variability in performance. Processing the data as a whole allowed identifying the events, procedures or methods that revealed to be more problematic for the users.

4 Results and Discussion

4.1 SINGRAR Usability Analysis Using the SUMI Method

SUMI method was applied to 13 users. All users answered the questionnaires correctly and their opinions were used to generate the results presented in Figure 2, in terms of Median and Upper and Lower Confidence Limit, for each of the five dimensions of usability (efficiency, affect, helpfulness, controllability and learnability). These results are synthesized by the Global usability assessment.

The analysis of the results allows concluding that users have a positive opinion about the system, i.e. equal to or greater than the commercial standard (the reference level 50), with some degree of dispersion in all dimensions of usability. The Global assessment with a value 60 and a small standard deviation indicates that the EMS is a software with high usability, better than the standard. Therefore, users were satisfied with the system, and to improve it only ad hoc corrections were needed. With the exception of items related to Control that were assessed as medium (controllability = 50), all items were assessed above the reference standard. The EMS is perceived by users as being very useful (helpfulness = 60), satisfactory (affect = 56), efficient (efficiency = 55) and relatively easy to learn (learnability = 54). The fact that the group of users who responded to the questionnaire is sufficiently large ensures that the analysis results are relevant.

Besides the general evaluation of the system's data, it was also performed an Item Consensual Analysis (ICA). The results related with 7 items (items 4, 6, 14, 22, 27, 29 and 41) of the SUMI questionnaire departed significantly from the pattern of response expected on the basis of the SUMI standardization database. In this set of items, four reflected a positive perception by users (items 6, 22, 27 and 41) and three reflected a negative perception (items 4, 14 and 29). The last three items were the ones that deserved special attention by the analysts, together with the software development team.

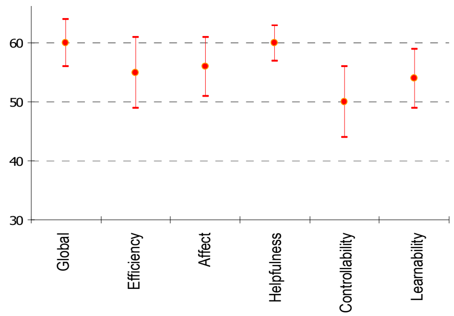


Fig. 2. SUMI Global Usability evaluation

The most important issues relate with the perception on the speed of the software (item 29) and to unexpected stops of the software (item 4). These issues were consistent with the concerns previously expressed by the users to the development team, and they were the main reason for the study. In fact, the main problems reported by users were related to the speed of data entry and to episodes of the software freezing.

Item 14 (related to the controllability of the system) was another area that deserved attention. Item 14 statement reads “I feel safer if I use only a few familiar commands or operations”. The results most probably reflected a lack of training on how the software works, and the recommendation of the study regarding this issue was that all operators should receive training, not to feel insecure when using the system.

The results for all other items addressed by the ICA indicate that the observed values did not differ substantially from the expected values, and the usability of the software components was not very different from the usability characteristics recorded in the SUMI database. There were several items where the software analyzed ranked better than standard, for example, items 12, 13, 24, 37 or 48. However, an example of items that were rated poorly is item 46, which is also an indicator of lack of knowledge about the system, which reinforces the previous recommendation on the need to increase education and training.

The use of the SUMI method offered a very good perspective about the level of quality of the EMS usability, and pointed to the need of implementing some minor modifications in the system, particularly in the domain of system control. In fact, this analysis together with a detailed interface and functionality analysis, allowed the identification of specific areas for improvement by the development team, that are discussed below. After these adjustments some gains in terms of system’s efficiency and effectiveness were obtained, which were evaluated and validated based on the objective analysis that is presented at the end of this section.

4.2 Dynamic Analysis of the Application

In order to analyze the aspects regarding the effectiveness and efficiency of EMS use, several sessions of objective data collection were conducted.

As referred, the measurements taken were obtained primarily from processing the Event Log files exported from the EMS at the end of each session. The analysis of these logs allowed, for example, obtaining data about the time spent to perform each task, and detecting the errors in inputting the events defined in the script. The analysis was complemented by the visualization of the videos recorded, enabling the review of the circumstances in which users performed the tasks, their comments and attitudes.

The processing and analysis of data collected in each individual session allowed understanding the circumstances that led to the specific results, in terms of time spent to perform the tasks and number of errors. This analysis allowed isolating aspects of the application, of the procedures associated with data entry and of the operation methods that proved problematic, or that were good practices to adopt. For instance it was possible to find that the way information was reported to users for data entry may significantly affect the amount of time spent to input it in the EMS and the number of errors, therefore affecting users' performance. Figure 3 illustrates this situation depicting the time spend to input data regarding a series of events. When the data is uniquely coded the time spent is consistently small. When the information is passed in a descriptive format the time required to process it presented high variability. This issue will be further discussed later in this sub-section.

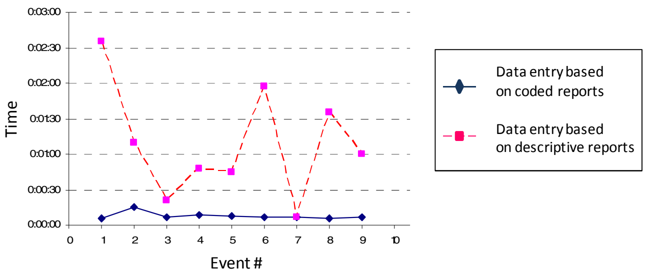
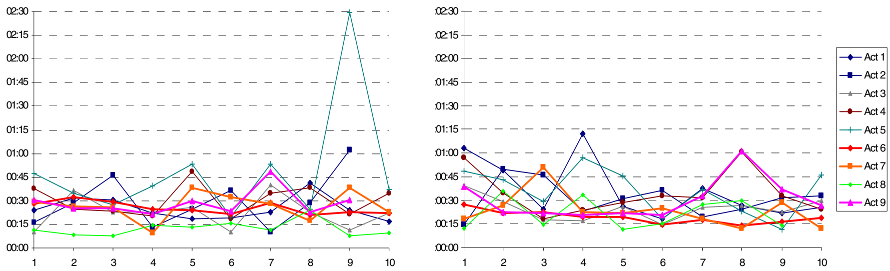


Fig. 3. Differences in data entry performance depending on the reporting procedure

On a second step the data were processed as a whole in order to identify how the events, procedures or methods imposed by the script affected the generality of the users group. This aggregation of individual data made possible identifying whether there were trends or variability in performance.

The operators engaged in the study presented different levels of experience, so that it was possible to evaluate if this factor affected in any way the user performance. The results demonstrated that experience was not significant in terms of the proficiency in data entry. It was observed that some of the users that were supposed to be more experienced presented levels of performance worse than the inexperienced users. The causes identified for this finding related mainly with the adoption of deficient procedures for using the system. Figure 4 illustrates the level of performance that could be reached by inexperienced users after the introduction of some improvements in the software and in the data input procedures, by comparing it with the output of the experienced user with best performance obtained in the tests done before improvements.



(a) Experienced user performance before improvements in the EMS

(b) Inexperienced user performance after improvements in the EMS

Fig. 4. Comparison of performance in data entry before and after improvements in usability

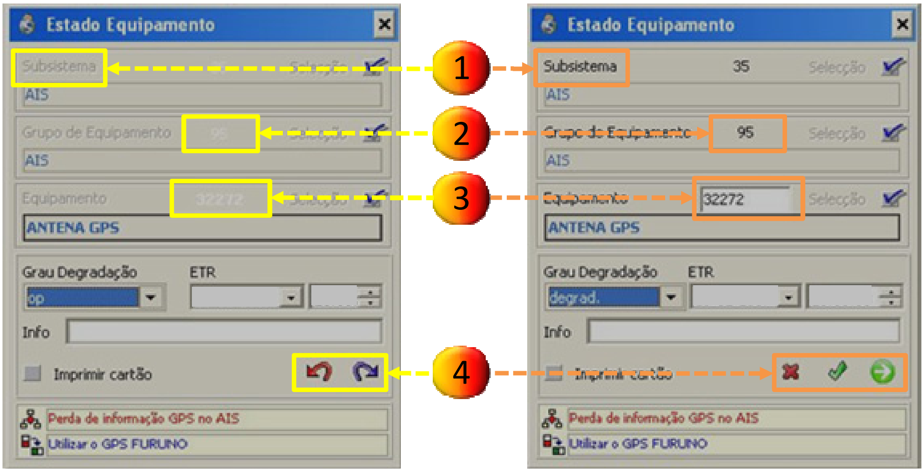
4.3 Analysis of Interfaces and User Interaction

The general arrangement of the interfaces was assessed considering the specificity of the context of use. Despite some complexity of the interface, the system was designed for a very specific objective and field of application. It was found that, in general, users did not have problems in accessing and using the features they needed.

The analysis of the interfaces was focused on assessing the graphical user interface (GUI), considering factors such as, the standardization of symbols and methods of accessing system functionalities, the type and size of lettering, and color contrast between letters and background. Figure 5 illustrates the before and after of a dialog box used to insert data in the EMS, which is representative of the type of intervention done. The dialog box shown on the left side depicts the design that was found in the beginning of the analysis where some deficiencies were identified. The analysts discussed the findings with the development team and some details were modified to improve interaction. The modified dialog box is depicted in the right side of the figure. Four modifications are highlighted in the figure, which will be discussed below as examples of the intervention that resulted from the usability analysis.

Considering the image contrast there were screens lacking adequate contrast between the background and the label letters, making it difficult to read, for example, white letters over a gray background (see 1, 2 and 3 in Figure 5). The symbology adopted failed to adhere consistently to common standards. Sometimes the icons chosen might lead users to make mistakes or were hard to relate to the functions they were associated with (see 4 in Figure 5). Similarly, the use of captions in the buttons was not consistent, since the terminology was not always the same. It was further observed that the operators tended to assume the existence and tried to use some mouse actions common in Windows environment (e.g., double-click), which were not always programmed, reducing efficiency and user satisfaction. The use of standard mouse forms of interaction such as double-click was recommended.

For experienced users the fastest way to access the features of a program is through the keyboard, for example using shortcut keys. This option was not systematically considered in the interface design, which limited the user's efficiency, since operation often required to take the hand out of the keyboard and to use the mouse to position the cursor over a button and to click. The use of shortcut keys was recommended.



(a) Initial GUI (b) GUI after usability analysis

Fig. 5. Examples of interface characteristics that were improved after the usability analysis

Data entry was often followed by the clicking of a command button (e.g., "OK"). Setting this button as the default button, improves user performance since it allows the operator to hit the "Enter" key in the keyboard, without having to handle the mouse. The use of default buttons was recommended for non-critical operations.

In order to expedite the selection of equipment for data entry, the "Equipment Code" (illustrated as 3 in Figure 5) became editable using a text box that allows writing directly the code of equipment. This solution avoided the need to manipulate other types of interfaces to select the equipment, saving significant amounts of time.

After finishing the entry of data regarding one event the software always closed the window presented in Figure 5. However this form is used repeatedly to input data. A new option was offered to users, which was to save the changes resulting from the entry of data of one event and to proceed with the introduction of data from a new event, without closing the window (see new icon in 4 of Figure 5). Besides this new feature, the cursor was also positioned automatically in the text box where the operator usually starts writing, thus avoiding the need to use the mouse to position the cursor. The implementation of these minor modifications significantly improved the efficiency in the entry of data of multiple events.

One issue that also had big influence on the systems effectiveness related with the performance of the search tool that was already available for finding equipment based on a textual descriptor. Although the algorithm worked properly, the Knowledge Base was not designed in a way that supported the system to recognize alternative acronyms (i.e., synonyms), making it difficult to find equipment that could be referred in the reports using multiple names or abbreviations. This limitation was noticed and the EMS Knowledge Base was updated, resulting in a major reduction of the time spent by users trying to locate equipment based on an equipment description.

5 Conclusions

Usability is a critical aspect to consider in the development cycle of software applications, and for this purpose, user-centered design and usability testing must be adopted, together with a particular care regarding the context of use of the software.

After performing the SINGRAR usability study it was possible to conclude that it offers a usability level considerably higher than commercial standard software. Users have a very favorable overall perception about the system.

The areas of concern identified by the SUMI method were analyzed and recommendations were produced to improve the interface design and speed up the data input process. This intervention originated gains in interface quality, and in operators' effectiveness and efficiency. The close cooperation between the analysts and the software development team was very fruitful in terms of identifying usability improvement areas and solutions.

Acknowledgements: This work was funded by National Funds through FCT - Fundação para a Ciência e a Tecnologia under the scope of project PEst-OE/EEI/UI0066.

References

1. Simões-Marques, M.J., Pires, F.J.: SINGRAR - A fuzzy distributed expert system to support Command and Control activities in Naval environment. European Journal of Operational Research 145(2), 343–362 (2003)
2. Nielsen, J.: Usability Engineering. Academic Press (1993)
3. ISO9241, Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on Usability. International Organization for Standardization (1998)
4. Simões-Marques, M., Nunes, I.L.: Usability of Interfaces. In: Nunes, I.L. (ed.) Ergonomics - A Systems Approach, pp. 155–170. InTech (2012), doi:10.5772/37299
5. Kirakowski, J.: The Use of Questionnaire Methods for Usability Assessment (1994), `http://sumi.ucc.ie/sumipapp.html` (accessed December 2012)
6. Nunes, I.L., Kirakowski, J.: Usabilidade de interfaces – versão Portuguesa do Software Usability Measurement Inventory (SUMI) [Interfaces Usability – Portuguese version of the Software Usability Measurement Inventory (SUMI)]. In: Arezes, P., Baptista, J.S., Barroso, M.P., et al. (eds.) Occupational Safety and Hygiene (SHO 2010), SPOSHO, Guimarães - Portugal, pp. 389–393 (2010)