SOFTWARE USABILITY EVALUATION
An Empirical Study

Tereza G. Kirner and Alessandra V. Saraiva
Graduate Program in Computer Science - Methodist University of Piracicaba
Rodovia do Açúcar, Km 156 - 13400-911, Piracicaba , SP - Brasil

Keywords: Software usability; Requirements usability evaluation; Empirical study; Goal-Question- Metrics.

Abstract: This article presents an empirical study performed to evaluate the usability of a software applied to the agri-livestock area. The evaluation plan was prepared with basis on the Goal-Question-Metrics paradigm. The research was performed in a government department of São Paulo state, in Brazil, and the subjects were professionals that give assistance to small rural properties in the planning, execution and control of agri-livestock activities, which can be supported by software systems. Usability is concerned with the suitability of the software for its users, defined in this work through the following attributes: easiness of understanding, easiness of learning, operationability, software attractiveness and user satisfaction, and usefulness and accomplishment of the goals. The preparation and execution of the empirical study are described and the data analysis and conclusions are presented. The obtained results indicate a satisfactory level of usability for the considered software. Besides evaluating the software, the study aim to contribute to the detailing of a process, based on GQM, to perform usability evaluations. The work also represents a contribution to the software quality improvement, primarily to those systems applied to agri-livestock tasks.

1 INTRODUCTION

Usability is a quality requirement of software products, which has been growing in importance mainly due to the contribution that it brings to the software improvement. Software with acceptable usability leads to positive impacts, such as: decrease in time and learning cost with the use of the system; easiness of transition to new system versions; users´ better performance and productivity; increase in the quality of tasks and decrease of mistakes due to data input provided by the users; etc. Moreover, several studies have already demonstrated that 80% of the total maintenance costs are related to problems that the users face on "what" the system makes and not with technical bugs. Of these problems, 64% are directly related to usability issues (Seffah, 2003).

Usability has been associated to attributes such as: effective use of the software; easiness of learning to use the software and exploring all its potentiality; reaching of the users' goals, through the software, in supported tasks; software attractiveness and users' satisfaction; conformity with specific patterns; and adaptation to the users' experience (Avouris, 2001; Bevan, 1991, 1995; Nielsen, 1993).

Studies have pointed out the need of methods that effectively help to identify and specify the users' needs, and to test and validate prototypes and software products with their final users (Seffah, 2003). In Brazil, particularly in the agri-livestock area, there are very few initiatives to identify and improve the quality of software products, fact that has been contributing to some retraction in the agro-informatics market (Cócaro, 2005).

This work presents an empirical study on the evaluation of the usability of a software applied to the agri-livestock area, aiming to contribute to the continuous improvement of software products in this area. At the same time, it intends to define a plan of usability evaluation together with a process of conducting empirical studies on this theme.

Section 2 summarizes some empirical studies on software usability, which were relevant for the present work. Section 3 presents the concepts in which the study was based on, including an explanation of usability and Goal-Question-Metrics (GQM). Section 4 presents the plan of usability evaluation, based on the GQM approach. Section 5 presents the definition of the empirical study, describing its accomplishment. The analysis of the

G. Kirner T. and V. Saraiva A. (2007).
SOFTWARE USABILITY EVALUATION - An Empirical Study.
In *Proceedings of the Ninth International Conference on Enterprise Information Systems - ISAS*, pages 459-465
DOI: 10.5220/0002352004590465

results is presented in section 6, and the conclusions are presented in section 7.

2 EMPIRICAL STUDIES ON SOFTWARE USABILITY

Among recent works reported in the literature, which provided a conceptual basis for the present work, studies to test the usability of electronic commerce systems (Rosa, 2004), education systems (Granic, 2004), and software packages (Dag, 2001) are highlighted. Moreover, a research about evaluation of agricultural software in Brazil, reported by Cócaro (2005), is indicated. These studies were reviewed to obtain a basis for the definition of usability and to determine the process of conducting the empirical study.

Rosa (2004) presents an empirical study on the usability of application forms in electronic commerce systems. The usability was defined through a series of aspects, such as: filling of the fields of the form, legibility, correction of filling mistakes, clarity of the instructions, explanatory and error messages, layout and organization of the form. To perform the study, two electronic commerce systems for book and CD sale were selected. Six subjects, from 20 to 40 years, who had already purchased on-line, participated of the study, responding to non-structured interviews. A qualitative analysis of the obtained data was performed, with the preparation of tables and calculation of percentages.

Granic (2004) presents a methodology to evaluate the usability of educational systems of ITS (Intelligent Tutoring Systems). The usability was identified through aspects related to the easiness of using the software, easiness of learning the software, and user satisfaction, such as: language and terms used in the software human-computer interfaces, supplied explanations, usefulness of the software, timing response, navigation through screens, and satisfaction provided by the accomplishment of the tasks. As part of the study, an ITS was evaluated by 5 usability experts. They answered a questionnaire with 10 questions, using a Lickert scale of seven points (Pereira, 2004), ranging from "I disagree" to "I agree". Data were tabulated, percentages where calculated regarding the answers, and the results were discussed.

Dag (2001) presents a case study for evaluating the usability of a software package produced by a Sweden software factory. The study used two usability evaluation methods: a questionnaire to allow a quantitative analysis, and a heuristic test, which allowed a qualitative analysis.

- For the questionnaire, it was used the SUMI (Software Usability Measurement Inventory) model (SUMI, 2006), which includes a standard framework for identifying usability issues. The questionnaire was sent to 90 potential final users of the considered software in Europe, and the results were statistically analyzed.
- The heuristic evaluation presupposes the experts' participation. In the study, 12 experts with knowledge on human-computer interfaces and on the software domain have participated. As result, it was obtained a list of usability problems, which should be corrected in order to produce an improved software version.

In the agri-livestock area, there are few experiences for obtaining and improving the quality of software products. This fact has been contributing for a retraction in the agro-informatics market (Cócaro, 2005). A recent study about evaluation of the quality of an agricultural software (Cócaro, 2005) focuses on the evaluation of two usability attributes: use easiness (including customization easiness and user's manual) and operational easiness (including simplicity to register operations, easiness of understanding the results, easiness of consulting the data and easiness of modifying the data). The evaluation methodology included the use of an evaluation record containing questions, which measure the quality attributes. Two specialists of the agri-livestock area answered the research, attributing concepts to the questions, followed by a discussion to justify the evaluation. The obtained results were organized with basis on the measured attributes.

The analysis of the empirical studies above indicated, suggested the need to invest in some aspects, such as:

- Definition of the usability requirement, aiming to reflect the real demands of specific end-user software;
- Preparation of the usability evaluation plan for the software;
- Definition of the process for performing the empirical study, including the definition of the objectives, the definition of the sample, the preparation of the instruments of data collection, the survey and preparation of the data, and the analysis of the results and identification of the conclusions.

The aspects highlighted above were focused on the present empirical study, described in the next sections.

3 CONCEPTUAL FOUNDATIONS

3.1 Software Usability

Usability is a software quality requirement that comprises aspects related to the efficient and effective use of the software. Several definitions are reported in the literature, including attributes such as: easiness of understanding the software functioning, easiness of learning the software, accomplishment of objectives and tasks, attractiveness of the software for the users, the users' satisfaction with the use of the software, adequacy with specific patterns, adaptation to the users' experience, etc. (Avouris, 2001; Bevan, 1991, 1995; Nielsen, 1993).

In the present work, the software usability was defined through the following attributes: easiness of understanding, easiness of learning, operationability, attractiveness and satisfaction of the user, and usefulness and accomplishment of goals. These attributes are defined as follows.

- Easiness of Understanding. Capacity of the software of enabling to the user the understanding for an appropriate use.
- Easiness of Learning. Allows the user to be able to explore the software and to accomplish his/her tasks without difficulties.
- Operationability. The software allows the user to operate and control it appropriately.
- Attractiveness and User Satisfaction. Capacity of the software to be attractive to the user, making him/her satisfied with the system interaction.
- Usefulness and Accomplishment of the Goals. The user reaches the goals proposed by the software, through its use.

The users' performance and expectations are directly related with the software usability. According to Govindarajulu (2003), the users of a system include several people, such as operators, developers and professionals that, in some way, control an application. The understanding of the users' profile and their roles are important for a study on software usability.

The present study considered, as final users, the agricultural technicians (agricultural engineers and veterinary physicians) who worked at a Department of Rural Development of São Paulo State, in Brazil, who had experience in the agricultural area and basic notions of Informatics.

3.2 Goal-Question-Metrics

The Goal-Question-Metrics **(**GQM) approach aims to support the elaboration and implementation of quality evaluation programs of products and software processes. GQM is based on the principles of Goal-Oriented Measurement and software Quality Improvement Paradigm (Basili, 1994; Kirner, 1997, Shull, 2006).

The preparation of a software measurement plan, based on GQM, comprises the following steps:

- Identification of the goals. Each goal is defined through the object, purpose, quality focus, point of view, and environment related to the intended study.
- Definition of the questions. For each goal, a set of questions has to be defined. These should be compatible with the proposed goal and include the requirements and quality attributes that are being evaluated.
- Definition of the metrics. Metrics to quantify the proposed questions have to be defined. Each metric should be related to at least one question, and each question should have at least one associated metric.

In this work, the GQM approach was used to support the elaboration of the usability evaluation plan for the considered software.

4 USABILITY EVALUATION PLAN

The measured software was developed by a Brazilian government organ, which support the decision-making related to solve specific problems of small, familiar properties, aiming at greater profitability for the rural producers (Saraiva, 2006).

The usability evaluation plan of the agricultural software was prepared with basis on the GQM approach. The goal, questions and metrics are presented as follows.

A brainstorm session was performed to start the definition of the evaluation plan, with the participation of 10 students enrolled in a software engineering discipline of a master's degree program. This session was very useful and it indicated a series of important aspects related to software usability.

4.1 Definition of the Goal

According to the GQM approach, the goal of the intended evaluation, which was to measure the

usability of an agri-livestock, was detailed. The goal is presented in Table 1.

Table 1: Goal of the Usability Empirical Study.

Goal		
Object	to analyze	Agri-livestock software
Purpose	with the purpose of	Evaluation of the software
Quality Focus	regarding to	Usability, evaluated through the following attributes: - easiness of understanding; - easiness of learning; - operationability; - attractiveness and user satisfaction; - usefulness and accomplishment of goals.
Point of View	under the point of view of	User - agriculture technician
Environment	in the following context	Department of Rural Development

4.2 Definition of the Questions

After the definition of the goal, a series of guidelines with 10 tasks was defined, to guide the users to work with the software. Based on these tasks, the questions were defined, which the users should answer in order to identify the level of software usability. Table 2 presents the defined questions for each usability attribute.

4.3 Definition of the Metrics

According to the GQM approach, metrics for the proposed questions, listed in Table 2, were defined. These metrics were based on a five-point Likert Scale (Pereira, 2004), according to which the users were requested to inform his/her degree of agreement or disagreement with each question, that is: (1): Disagree Completely; (2): Disagree; (3) Neither Agree nor Disagree; (4): Agree; (5): Agree Completely.

5 EMPIRICAL STUDY

5.1 Objective of the Study and Definition of the Sample

The empirical study aimed to evaluate the usability of a software applied to the agriculture area, measured through the following attributes: easiness of understanding; easiness of learning; software operationability; software attractiveness and user satisfaction; and usefulness and accomplishment the goals.

The population of the empirical study was composed of the agricultural technicians of a Brazilian Department of Rural Development. After the population was established, the sampling technique adopted to choose the elements that would compose the sample was defined. Two aspects were relevant for defining the sampling: the fact that the participants have specific characteristics, such as, being technicians with higher education, and having basic notions of the use of computers and Informatics. In this way, it was decided to use a convenience sampling, according to which the subjects are chosen because they have characteristics that are appropriate to the study.

The sample was composed of 11 subjects, whose characteristics were according to those mentioned above. From them, 54,55% were agriculture engineers and 45,45% were veterinary physicians.

5.2 Instruments for Data Collection

The present study used three instruments for data collection, which were prepared based on the evaluation plan presented in the previous section. Those instruments, available in Saraiva (2006), are:

- Questionnaire for General Information Surveying. These questionnaires collected information on the profile of the participants, such as academic grades and level of experience.
- Guidelines for Using the Software. They included 10 topics to guide the subjects to perform tasks regarding the use of the software. Starting from the accomplishment of the defined tasks, the subjects would be able to evaluate the software usability.
- Questionnaire of Software Usability Evaluation. It was composed by 20 questions, to be answered with basis on the Likert scale. The evaluators should answer each question, underlining one of the presented values, which are: (1) Agree completely; (2) Agree; (3) Neither agree nor disagree; (4) Disagree; (5) Disagree completely.

To guarantee the reliability of the collected data and a greater index of answers, the questionnaires were applied through individual interviews, performed with the participants, in the facilities of the Department. Besides, a pilot test was performed to test the instruments of data survey before using them in the research. The test had the collaboration of a user of the agricultural area, with similar profile of the further evaluators.

Table 2: Questions for the Empirical Study.

Questions related to the Software Easiness of Understanding
- Q1. In the "Main Screen", the illustrations of buttons help the understanding of the software.
- Q2. In case there is the need to use the Option HELP, the user is able to use it.
- Q3. During the use of the software, the user notices the objective of the application.
- Q4. Menu dispositions are easily understood during the sequence of tasks.
Questions related to the Software Easiness of Learning
- Q5. The language of the software in "Portuguese" is considered a relevant factor for the learning of the software.
- Q6. The software is considered "Self-explanatory" in its functions.
- Q7. In the "Main Screen", the disposition of the illustrations of buttons helps the learning of the software.
- Q8. The use of the Menus provides the learning of the software.
Questions related to the Software Operationability
- Q9. The user considers that the navigation among screens enables the performance of operations.
- Q10. The icons are representative in the use of the software.
- Q11. The user can easily register the "property".
- Q12. The user can clearly interpret the graphic representation.
Questions related to the Software Attractiveness and Users' Satisfaction
- Q13. The illustrations of buttons are attractive.
- Q14. The user feels pleased when using the software.
- Q15. The use of this software can collaborate with the planning of agricultural activities.
- Q16. The fact that it is a free distribution software is an attractive reason for its use.
Questions related to the Software Usefulness and Accomplishment of Goals
- Q17. The software is an advantageous tool for agricultural development.
- Q18. The software usefulness was insufficient in some applications.
- Q19. The software fulfilled the needs of your business.
- Q20. You could reach your goals with the use of the software.

He was guided to identify ambiguous or confused aspects, in the three questionnaires. The identified aspects were discussed, analyzed and corrected.

6 ANALYSIS OF RESULTS

The analysis of the results followed guidelines supplied by statistics (Pereira, 2004). A suitable methodology was adopted for qualitative data, using descriptive techniques, Likert scale and the Cronbach's Alpha coefficient.

Data obtained through the Questionnaire of Software Usability Evaluation allowed the evaluation of the software usability regarding the considered attributes - easiness of understanding, easiness of learning, operationability, software attractiveness and user satisfaction, and usefulness and accomplishment of goals.

The totals (in percentages) of the values marked by the participants to each usability attribute are presented in Table 3. The answers, chosen by the participants according to the Lickert scale, presented the average of 4,2818. That means that the point of balance of the informed answers pointed out a relative agreement to the questions of the Questionnaire of Software Usability Evaluation.

Table 3: Totals (in %) of Values per Attribute.

attributes	Complet agree	Agree	Neither agree nor disagree	Dis-agree	Complet disagree
Easiness of Underst.	47,73%	52,27%	0%	0%	0%
Easiness of Learning	43,18%	50%	2,2%	4,55%	0%
Operationa-bility	40,91%	56,82%	0%	2,27%	0%
Attract. and User Satisfaction	52,27%	43,19%	4,55%	0%	0%

The results obtained for the usability attributes are presented in the next subsections.

6.1 Easiness of Understanding

The questions Q1 to Q4 of the Questionnaire of Software Usability Evaluation measured the software usability regarding to the "easiness of understanding" attribute.

For these questions, the answers "Neither Agree nor Disagree", "I Disagree", and "I Disagree Completely", had null result in the users' opinion. In the other way, the answers "I Agree Completely"

and “I Agree" had a value of 47,73% and 52,27%, respectively. This shows that the users, in a general way, agree that the considered agricultural software is easy to understand.

6.2 Easiness of Learning

The questions Q5 to Q8 of the Questionnaire of Software Usability Evaluation measured the software usability regarding to the “easiness of learning” attribute.

For these questions, the answer was only given once, that is, a single user was positioned as indifferent to the easiness of learning characteristic provided by the software, which represents 2,27% of the responses. As for “I Disagree", two users chose this answer, representing 4,55% of the total percentile. The answer "I Agree Completely" represented 43,18% of the total percentile. Moreover, the most significant fact is that 50% of the users that participated in the evaluation "Agree" that the evaluated agricultural software is easy of learning, during its use.

6.3 Operationability

The questions Q9 to Q12 of the Questionnaire of Software Usability Evaluation measured the software usability regarding to the “operationability” attribute.

For these questions, the answer "I Disagree" was chosen only by one user, which represents 2,27% of the total percentage. The "Neither Agree nor Disagree" and “I Disagree Completely" answers were not marked. The answer "I Agree Completely" had a percentage of 40,91% according the users' evaluation, and the answer " I Agree" represented 56,82% of the marked answers. Such results suggested that the operationability of the evaluated software is an attribute noticed by the users while they are using it.

6.4 Attractiveness and User Satisfaction

The questions Q13 to Q16 of the Questionnaire of Software Usability Evaluation measured the software usability regarding to the “attractiveness and the user satisfaction” attribute.

For these questions, only two users answered that are indifferent (“Neither Agree nor Disagree”) to the software attractiveness and/or the user satisfaction, which represents only 4,55% of the total percentage. From the remaining users, 43,19% "Agree" and 52,27% "Agree Completely" that the attractiveness is present in the focused agricultural software and that there is user satisfaction in using the software.

6.5 Usefulness and Accomplishment of Goals

The questions Q17 to Q20 of the Questionnaire of Software Usability Evaluation measured the software usability regarding to the “usefulness and achievement of the goals” attribute.

Concerning this attribute, a larger degree of diversified answers was observed, since all the answer categories were selected. The answers "I Disagree Completely" presents 4,55%, the "Neither Agree nor Disagree” option shows 9,19%, "I Disagree" indicates 15,91%, and “I Agree Completely" presents 31,82%cof the total percentage. The most suitable option for the users, with 38,64 % of the answers, is "I Agree".

6.6 Analysis of Consistency and Reliability

The analysis of the intern consistence of the Questionnaire of Software Usability Evaluation was estimated through the Cronbach’s Alpha coefficient, which allowed an evaluation of the reliability of the research results (Pereira, 2004).

In the present study, the representativeness of Alpha Coefficient is 0.8743, suggesting that it is measuring 87,43% of the real impact. This value of Alpha is satisfactory in terms of reliability of the results, since it is very close to 100% of the evaluation efficiency. The indicators show that the questionnaire is concise and the questions regarding the usability attributes evaluated in the empirical study have coherence and authenticity.

7 CONCLUSION

This paper presented an empirical study about usability evaluation of a software for the agri-livestock area. The empirical study was carried out in a Brazilian department of supporting for agriculture and livestock in São Paulo state.

To guide the research, a plan of usability evaluation was elaborated, based on the GQM approach. This plan was useful for the complete understanding of the problem, enabling the clear accomplishment of the proposed goal for the study, the identification of the questions to be answered,

and the definition of the metrics related to the questions. The plan helped the definition and preparation of the research instruments, guaranteeing the creation of questionnaires that included the usability aspects.

This empirical study followed a process based on statistical techniques, which allowed a qualitative analysis of data, providing clear identification of the results. The use of Cronbach´s Alpha coefficient could determine the reliability of data and results, which were significant for the study.

In the area of agriculture, the study showed the usability of a software used to support small and medium rural properties. Thus, a contribution was made to the continuous improvement of software in this area. It is important to point out that the obtained results indicated a satisfactory level of usability for the considered software.

Besides evaluating the software, the study contributed to the detailing of a process, based on GQM, to perform usability evaluations. The work also represents a contribution to the software quality improvement, primarily to those systems applied to the agriculture.

Finally, it is important to highlight the need of investing in new research that could contribute to the results of this study. Therefore, such results could be extended to a broader spectrum of software, in the agricultural area as well as in other ones.

ACKNOWLEDGEMENTS

The authors would like to thank Idemauro de Lara and Maria Imaculada Montebelo, for the valuable support in the data analysis.

REFERENCES

Avouris, N.M., 2001. An introduction to software usability. In *Workshop on Software Usability – Proceedings of the 8 th Panhellenic Conference on Informatics*. v. 2, p. 514-522.

Basili, V.R.; Caldiera, G.; Rombach, H.D., 1994. Goal-Question-Metrics paradigm. *Encyclopedia of Software Engineering*, John Wiley & Sons. New York, vol. 1, p. 527-532.

Beavan N., Kirakowski, J.; Maissel, J., 1991. What is usability? In *Proceedings of the 4 th International Conference on Human-Computer Interaction*. Stuttgart, Germany.

Beavan N., 1995. Usability is quality of use. In *Proceedings of the 6th International Conference on Human-Computer Interaction*. Tokio, Japan.

Cocaro, H.; Lopes, M. A.; Campos, F.C., 2005. Quality of agri-livestock software: a case study. *Science and Agro-Technology*, Lavras, MG, Brazil, v. 29, n. 5, p. 1075-1082.

Dag, J.N et all., 2001. An industrial case study of usability evaluation. In *Proceedings of the 9th International Conference on Human-Computer Interaction*. Mahwah, NJ, p.425-429.

Granic, A.; Glavinic, V.; Stankov, S., 2004. Usability evaluation methodology for web-based educational systems. In *Proceedings of the 8th ERCIM Workshop*.

Govindarajulu, C., 2003. End Users: Who are they? *Communications of the ACM*, Vol. 46, N. 9, p. 152-159.

Kirner, T.G.; Abib, J.C., 1997. Use of the Goal-Question-Metrics in the evaluation of software quality requirements. In *Proceedings of the IV Brazilian Workshop on Software Quality*. Fortaleza, CE, Brazil, p.71-82.

Nielsen, J., 1993. *Usability Engineering*. Academic Press, Cambridge, MA.

Pereira, J.C., 2004. *Analysis of Qualitative Data*. São Paulo, University of São Paulo Press, 2004.

Rosa, J.G.; Moraes, A., 2004. Usability in electronic commerce forms. In *Congress of Ergonomy of the Countries of Portuguese Language*. Funchal, Island of Madeira. (in Portuguese).

Saraiva, A.V., 2006. *Use of the Goal-Question-Metrics in theElaboration and Execution of Usability Measurement Plans*. Master Thesis, Graduate Program in Computer Science, Methodist University of Piracicaba, Brazil (in Portuguese).

Seffah, A.; Benn, J., 2003. Making usability a respectable quality attribute in the engineering lifecycle. *IEEE Canadian Review*, Fall 2003, p. 25-28.

Shull, F.; Seaman, C.; Zelkowitz, M., 2006. Victor Basili's contributions to software quality. *IEEE Software*, Volume 23, Number 1, p. 16-18.

SUMI, 2006. *The SUMI Model*. Retrieved October 2006 from http://sumi.ucc.ie.