Quantitative Usability Measurement for Commercial Off-the-Shelf Products: the case of Core-Banking System

Lemlem Berhanu
Commercial Bank of Ethiopia, Addis Ababa, Ethiopia
lemruth@yahoo.com

Rahel Bekele
School of Information Science, Addis Ababa University, Ethiopia
rahtesf@yahoo.com

Abstract

This paper assesses how usability is perceived and considered as evaluation criterion at the time of purchasing COTS products. Additionally, how far a COTS product is usable for its users and the extent of user satisfaction is also evaluated. Further literature study, survey, and its findings are discussed. Finally, COTS Usability Questionnaire is developed to quantitatively measure usability for COTS products.

Keywords: Usability; Quantitative Usability Measurement; Core-Banking System; COTS Product; COTS Usability Questionnaire

1. Introduction

Banks and financial service organizations have long recognized the advantages of deploying technology to improve value, speed and, flexibility of their product offering to customers [1]. Software developed in-house or Commercial Off-the-Shelf (COTS) products are used to support their products and services.

Among COTS products, Core-Banking System (CBS) is used by commercial banks. CBS is a vital element that helps a bank to differentiate itself and allows offering many new technology driven channels to customers such as Automated Tellers Machines and Internet Banking [1]. In order to systematically evaluate, rank, and select a COTS product that best meets the bank requirements as presented by Ncube and Dean [2], different criteria like cost, usability, functionality, size, portability, supplier capability, etc. are used. The focus of this paper is on usability as evaluation criterion of COTS products at the time of purchasing.

2. Statement of the Problem

Usability is increasingly recognized as an important factor for interactive software systems like CBS [3]. Moreover, Ovaska [4] states that lack of usability evaluation of software packages thoroughly leads to acquisition of software with poor usability. Poor usability increases the time needed to get desired results, increases errors, and causes misunderstanding, even misjudged decisions [4].

The National Bank of Ethiopia is on the way of implementing National Payment System (NPS) which supports straight through processing between commercial banks which requires implementation or upgrading of CBS [5]. Furthermore, the cost of purchasing a CBS is a huge investment not only for commercial banks but even for the country as well [6].

Since usability is a quality that many products possess, but many lack due to historical, cultural, organizational, monetary, and other reasons [7], focusing on usability of COTS products is found to be important. Therefore, the aim of this paper is to quantitatively measure usability of COTS products.

3. Survey Results

A survey was undertaken in six selected commercial banks using questionnaire in order to identify the current practice of usability in procurement of COTS products. Moreover, usability testing was also conducted on selected CBS to assess the extent of user satisfaction and how easy a COTS product is for actual users. According to the study analysis

- Usability is defined as user friendliness, ease of use and quality of software by majority of the respondents, whereas others

misconceived it as an attribute of the interface.

- Usability is not considered as one of the evaluation criteria like functional requirements, non-functional requirements, vendor profile, product demonstration, reference check, and negotiations.
- Usability evaluation methods that are used are not appropriate to measure usability of COTS products.
- Usability testing proves the existence of poor efficiency, "difficult" tasks and poor level of satisfaction for COTS products. These confirm that the COTS product is not fairly usable for its real users.
- Challenges related with usability at time of COTS product procurement:
  - A requirement that was not well defined in the Request for Proposal (RFP).
  - Difficult to observe usability based on vendor's response from an RFP document. Due to this, product demonstration and reference check are used to cross check usability of the candidate CBS.
  - A problem of getting usability experts.

4. The Proposed Solution

A COTS Usability Questionnaire is developed to quantitatively measure usability of COTS products. The questionnaire contains forty items with a five-point Likert scale with anchors ranging from "Strongly Disagree" to "Strongly Agree" as shown in Table 1. The questionnaire items include factors such as learnability, flexibility, memorability, efficiency, effectiveness, and satisfaction.

Table 1: COTS Usability Questionnaire

No.	*Question*	*Strongly Disagree*	*Disagree*	*Undecided*	*Agree*	*Strongly Agree*
1.	My previous experiences of using other systems help me learn the system very easily.	O	O	O	O	O
2.	It is tiring to click multiple buttons to confirm an action.	O	O	O	O	O
3.	The system helped me to perform my assigned tasks accurately and timely.	O	O	O	O	O
4.	The system is pleasant to use.	O	O	O	O	O
5.	The error messages clearly tell me how to fix problems.	O	O	O	O	O
6.	Messages which appear on the screen are sometimes confusing.	O	O	O	O	O
7.	Using the system, tasks can be performed in a straight forward manner.	O	O	O	O	O
8.	Working with the system is satisfying.	O	O	O	O	O
9.	It takes too long to learn commands/menus to operate the system.	O	O	O	O	O
10.	After a period of not using the system, I can easily use it.	O	O	O	O	O
11.	Instructions and prompts are helpful and clear.	O	O	O	O	O
12.	I feel in command of the system when I use it.	O	O	O	O	O
13.	Learning to operate this system initially is full of problems.	O	O	O	O	O
14.	The system provides me options to recover from mistakes quickly and easily.	O	O	O	O	O
15.	I sometimes do not know what to do next with the system.	O	O	O	O	O
16.	I would not like to use this system.	O	O	O	O	O
17.	The commands relevant to a set of tasks are easy to learn.	O	O	O	O	O

No.	*Question*	*Strongly Disagree*	*Disagree*	*Undecided*	*Agree*	*Strongly Agree*
18.	The system makes me recognize how to perform a task instead of remembering.	O	O	O	O	O
19.	The help information given by the system is not very helpful.	O	O	O	O	O
20.	Using the system is frustrating.	O	O	O	O	O
21.	I have to look for assistance when I use this system.	O	O	O	O	O
22.	The flow of the system reflects how the user thinks of the work flow.	O	O	O	O	O
23.	The system gives me control over my activities.	O	O	O	O	O
24.	I would recommend this system in other organizations.	O	O	O	O	O
25.	I feel that a lot of training will be needed for beginners to make use of the system.	O	O	O	O	O
26.	I can easily customize/adapt the system.	O	O	O	O	O
27.	The reports generated are not often reliable and accurate.	O	O	O	O	O
28.	I found the system useful in my work.	O	O	O	O	O
29.	I have easily learnt the steps needed to accomplish what I want to do.	O	O	O	O	O
30.	I can easily and successfully find the help options.	O	O	O	O	O
31.	It is relatively easy to move from one part of a task to another.	O	O	O	O	O
32.	I am willing to teach others on how the system works.	O	O	O	O	O
33.	It is easy to make the system do exactly what you want.	O	O	O	O	O
34.	I can always make safe exits from the system whenever I want too.	O	O	O	O	O
35.	The response time of the system is the cause of my frustration.	O	O	O	O	O
36.	The system does everything I would expect it to do	O	O	O	O	O
37.	The logical grouping and ordering of menu options helped me learn the system.	O	O	O	O	O
38.	I feel safer if I use only a few familiar commands or operations.	O	O	O	O	O
39.	Instructions for commands or choice are clear.	O	O	O	O	O
40.	I like the order of presentation of interfaces.	O	O	O	O	O

Learnability means the system should be easy to learn so that users can start getting some work done. Flexibility and memorability indicate the system is easy to remember and a user is able to use the system after some period of not using it, without additional learning. Moreover, users are able to customize aspects of the system. Efficiency and effectiveness mean once the user has learned the system, a high level of productivity is achieved by accurately completing specified tasks with expected time. Satisfaction shows how users are subjectively satisfied by using the system.

The most important and fundamental characteristics of any measurement procedure, reliability and validity, were assessed by a survey collected from 200 COTS product users. The analysis showed Cronbach's Alpha of about 0.9 that is good internal consistency. All factors regarding learnability, efficiency and effectiveness as well as satisfaction have a reliability result which shows good internal consistency. But factors grouped as flexibility and memorability resulted in a reliability index less than the minimum suggested value (i.e., 0.7). It also has good construct validity.

In order to quantitatively measure usability of COTS products, first representing the responses given to each item score contribution with a range from 0 (i.e., "Strongly Disagree") to 4 ("Strongly

Agree") is required. For negatively worded items, the contribution ranges vice versa that is 4 as "Strongly Disagree" to 0 as "Strongly Agree". Evaluated usability is a product of total average value, which is the sum of all respondents' total score divided by total number of participants by item value (i.e., 0.625).

As pointed by Bangor *et al*. [8], using a Standard Letter Grade Scale, products that scored in the 90s were exceptional, products that scored in the 80s were good, and products that scored in the 70s were acceptable. Anything below a 70 had usability issues that are causes for concern. Based on this, interpreting quantified usability of COTS products makes it easy to understand the implication by many people and it facilitates communication of results.

5. Relate Work

Though usability is becoming a more and more important software criterion, available usability measurement methods are either difficult to apply, or overly dependent upon evaluators' expertise.

Lin *et al*. [9] identified eight human factor considerations (i.e., compatibility, consistency, flexibility, learnability, minimal action, minimal memory load, perceptual limitation, and user guidance) which are relevant to software usability along with the three stages of human information (Perceptual, Cognitive and Action) processing theory which formed the framework from which Purdue Usability Testing Questionnaire (PUTQ) was derived. PUTQ is intended mainly to measure users' satisfaction toward a Graphical User Interface (GUI). Moreover, PUTQ applicability for COTS products is limited since it doesn't take into account some usability factors such as effectiveness and efficiency.

Furthermore, Lewis [10] used a rank-based system when assessing competing products. The technique works with data collected during scenario-based usability studies where participants are asked to perform realistic tasks with products. This approach creates a rank score comprised of both users' objective performance measures and subjective assessment. Application of the method allows a single composite rank-average to be assigned to a product to represent its relative usability by allowing easy comparison of products. But the resulting metric only represents a relative comparison between like-products with similar tasks. It does not result in an absolute measure of usability that can be compared across products of different task-sets.

The rank-based method is meant for software development and did not consider COTS products from the procurement perspective. Furthermore, it doesn't give absolute result for usability of the software under testing, which makes it difficult to adopt the method for COTS products usability testing. This is due to the fact that COTS products with poor usability might get the first rank when compared to other candidate COTS products. This misleads the evaluation committee members to make a wrong decision to purchase the software.

Several usability measurement models have been proposed to measure and report usability of a general software product. Sauro and Kindlund [11] proposed a quantitative model which simplifies usability dimensions into a single, standardized and Summated Usability Metric (SUM). In SUM, task completion, error counts, task times, and satisfaction scores metrics are used in order to represent effectiveness, efficiency, and satisfaction usability dimensions. A summative evaluation is used to quantitatively assess the "before and after" impact of design changes which provides an idea of how usable a task or product is without having to reference historical data.

On the other hand, Seffah *et al*. [3] come up with Quality in Use Integrated Measurement (QUIM), a consolidated hierarchical model for usability measurement which decomposes usability into factors, criteria, and metrics. The main application for QUIM is to provide a consistent framework and repository for usability factors, criteria, and metrics for educational and research purposes. The goal of developing QUIM is to reduce software development risks in a less expensive approach.

Most of the tools that exist for assessing software usability are usually proprietary and may only be available for a fee. In addition, these tools were intended to be used at software development lifecycle rather than for procurers of software at final software product level.

6. Conclusion and Future Work

In this paper, COTS Usability Questionnaire is presented to measure usability of COTS products. It is useful for commercial banks and other COTS purchasing organizations to make appropriate evaluation in terms of COTS product's usability at the time of purchase. The research provides information on how usability is perceived and gives insight for decision makers, how far the current CBS is usable or not for its intended users. Furthermore, it pinpoints the banking community to consider usability as one of the requirements and criteria at the time of purchasing COTS products.

Further task is required to amend the questionnaire so that it can yield better reliability index for the usability factors flexibility and memorability. Moreover, the proposed solution quantifies usability of COTS products by giving equal weight to usability factors. It is recommended to research about weighting the usability factors and automation of the COTS Usability Questionnaire.

Even though the majority of commonly used questionnaires were initially developed in the English language, as pointed in [12], different types of questionnaires have been widely translated into other languages. Future work in this study is recommended to translate the COTS Usability Questionnaire into local languages like Amharic.

References

[1] Infrasoft Technologies, "End-to-End Core Banking Solution for Competitive Advantage," White Paper, 2009, http:// cache-www.intel.com/cd/00/00/19/30/ 193047_193047.pdf, last accessed on December 10, 2011.

[2] C. Ncube and J. Dean, "The Limitations of Current Decision-Making Techniques in the Procurement of COTS Software Components," *in Proceedings of the 1st International Conference*, 2002, pp. 176-178.

[3] A. Seffah, M. Donyaee, R. B. Kline, and H. K. Padda, "Usability Measurement: A Roadmap for a Consolidated Model," *Software Quality Journal*, Vol. 14, No. 2, 2006, pp. 159–178.

[4] S. Ovaska, "Usability as a Goal for the Design of Computer Systems," *Scandinavian Journal of Information Systems,* Vol. 3, 1991, pp. 47-62.

[5] J. Woltjer, L. Elmasry, and V. Babin, "Modernization of the National Payment Systems in Ethiopia, Part 3, Vision and Strategic Framework," Version 6, 2009.

[6] IBM, "Commercial Bank of Ethiopia Signs Strategic Deal with IBM to Modernize Core Banking Systems and Support Launch of 500 New Branches," 2011, http://www-03.ibm.com/press/us/en/pressrelease/34913.wss, last accessed on December 01, 2011.

[7] J. Rubin and D. Chisnell, *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*, 2nd ed., Wiley Publishing, 2008.

[8] A. Bangor, P. Kortum, and J. Miller, "Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale," *Journal of Usability Studies*, Vol. 4, Issue 3, 2009, pp. 114-123.

[9] H. X. Lin., Y.Y. Choong, and G. Salvendy, "A Proposed Index of Usability: A Method for Comparing the Relative Usability of Different Software Systems," *Behaviour and Information Technology*, Vol.16, 1997, pp. 267-277.

[10] J. R. Lewis, "A Rank-Based Method for the Usability Comparison of Competing Products," in *Proceedings of the Human Factors Society 35th Annual Meeting*, San Francisco, Human Factors Society, 1991.

[11] J. Sauro and E. Kindlund, "A Method to Standardize Usability Metrics into a Single Score", in *Proceedings of the Conference in Human Factors in Computing Systems (CHI 2005),* Portland, 2005.

[12] C. W. Sun, "Questionnaire Translation and Psychometric Properties Evaluation," SEGi University College, Vol. 2, No. 2, 2009, pp. 45-51.