World Academy of Science, Engineering and Technology
International Journal of Industrial and Systems Engineering
Vol:3, No:7, 2009

A Usability Testing Approach to Evaluate User-Interfaces in Business Administration

Salaheddin Odeh and Ibrahim O. Adwan

Open Science Index, Industrial and Systems Engineering Vol:3, No:7, 2009 publications.waset.org/1414.pdf

Abstract**—**This interdisciplinary study is an investigation to evaluate user-interfaces in business administration. The study is going to be implemented on two computerized business administration systems with two distinctive user-interfaces, so that differences between the two systems can be determined. Both systems, a commercial and a prototype developed for the purpose of this study, deal with ordering of supplies, tendering procedures, issuing purchase orders, controlling the movement of the stocks against their actual balances on the shelves and editing them on their tabulations. In the second suggested system, modern computer graphics and multimedia issues were taken into consideration to cover the drawbacks of the first system. To highlight differences between the two investigated systems regarding some chosen standard quality criteria, the study employs various statistical techniques and methods to evaluate the users' interaction with both systems. The study variables are divided into two divisions: independent representing the interfaces of the two systems, and dependent embracing efficiency, effectiveness, satisfaction, error rate etc.

Keywords**—**Evaluation and usability testing, software prototyping, statistical methods, user-interface design.

I. INTRODUCTION

ONE of the most important concerns in modern administration is how to utilize the available recourses with effectiveness, efficiency and satisfaction. Socio-technical systems that offer a basis for collaborative environments [1] can give a theoretical frame-work for the design of guidelines within the context of group work-oriented technology. The ease or comfort during the usage of computerized administration systems is mainly determined by characteristics of the software product itself, such as the user-interface. As usability is concerned with the usage of software systems generally, it forms an important pillar for determining and optimizing of product quality characteristics. Usability takes into consideration the type of the users, the tasks to be carried out, as well as the physical and social aspects that can be related to the usage of the software products.

The interaction between users and interactive software system can be evaluated since the users can access the system through keyboards or pointing devices, and can get a continuous feedback through displays on screen. When evaluating a product or series of products, one may make a product-against-product comparison, or compare each product with standardization database, to see how the product that is being rated can be compared against an average state-of-the-market profile. Here, we will hold a comparison between an employed system and a prototype, which is not only designed for the purpose of evaluation, but also for building a new product with enhanced user-interface and optimized functionalities as well. Back to socio-technical systems, these may be conceived as consisting of three non-trivially interlinked systems: individual persons, organization and technology. A socio-technical system does not only include the technical system, but it also has operational processes people using or interacting with it. Moreover, these systems are governed by organizational policies and rules. Whenever one of the sub-systems changed, the others will be affected as the case in usually complex fashions. Thus, the design of such systems has to be deliberated jointly with the socio-technical system design as a whole.

To strengthen the inter-relations between individuals and organizations, tasks should be designed in a way to enable the working person to make the full range of his/her skills, and at the same time to require continuous development of these skills and experience. If individual learning process is to have any impact on the organizational learning, the continuous redesign of work processes must be a part of the workers' tasks [2]. In software engineering, usability is a criterion consisting of several factors such as: ease to use, orientation, utility, etc [3]. The main design rule for the relation between organization and technology is that, technology should fit the organizational structure of the company, e.g. decentralization work processes in semi-autonomous groups, and should never, by its own interval structures, impose certain organizational structure on the company.

In the relation between individual and technology, the design of technical systems should, on the one hand, fit the skills, experiences and ways of thinking and acting of the workers, and, allow and foster learning by offering flexible modes of actions for a step-by-step increase in competence on the other. When planning to employ any software system, it is difficult to integrate heterogeneous information before processing it to the plan phase, which has two main problems for constructing information system for strategic planning: namely, information type or form to be used for strategic planning, cannot be defined beforehand, as well as the reconstruction of existing information systems is very expensive [4].

In this study, we will review ways of applying usability engineering for evaluating user interfaces in business administration by comparing two computerized systems

S. Odeh is with Department of Computer Engineering, Faculty of Engineering, Al-Quds University, Abu Dies Jerusalem, Palestine (corresponding author to provide phone: ++972-2- 279 7024; fax: ++972 -2- 279 7023; e-mail: sodeh@eng.alquds.edu).

I. O. Adwan is UNRWA staff member – Administration Department – Jerusalem, Palestine (corresponding author to provide phone: ++972-2- 297 5781/2; fax: ++972 -2- 297 2968; e-mail: i.othman@unrwa.org)

depending on Norman's model [3]. This model is suitable for the purpose due to the facts that it has its disseminated implementations in office software applications, internet and other administrative programs; it is a non-complicated model which gives a precise description of the user's reactions. It should be noted that usability tests for evaluating interactive software systems are beneficial and thus profitable, as after the well-known rule of thumb every $10 invested in usability will achieve a return of $100 [5].

This study is interdisciplinary as two researches from two different fields, namely business administration and computer engineering, were attempting together to evaluate graphical user-interfaces in business administration through having applied statistical methods to achieve, effective human-computer interaction, and to contribute to the usability engineering field in the e-business administration. Apart from the user-interfaces of the two investigated systems, the two business administration systems are actually the same since this information system follows the architectural design of a layered model of an information system (see Fig. 1), where the top layer backs supports the user-interface and the bottom layer the system database. The user communications layer handles all input and output from the user-interface, and the information retrieval layer includes application-specific logic for accessing and updating the database [1]. Due to the facts that the two systems have diverse user-interfaces, it was anticipated to achieve miscellaneous task results of the users interacting with both systems.

While the first system is commercial, the second is a prototype developed for the purpose of this investigation. The two systems enable us to carry out business-administrative tasks such as ordering of supplies, tendering procedures, issuing purchase orders, controlling the movement of the stocks against their actual balances on the shelves and editing them on their tabulations. When having developed the prototype, whose development and implementation followed the iterative design [6], several contemporary aspects of human-computer interaction enabled through computer graphics and multimedia were taken into account to bypass the drawbacks of the commercial system. The study takes into account the statistical issues in order to evaluate the users' interaction with the two systems, and consequently to highlight differences between them regarding diverse quality criteria. The study variables are divided into two divisions: independent representing the interfaces of the two systems, as well as dependent that include efficiency, effectiveness, satisfaction, error rate etc.

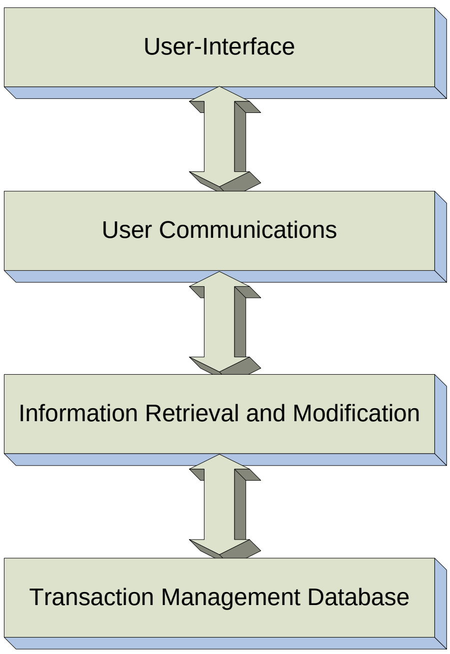

Fig. 1 A layered model of an information system [1]

There are many usability methods and techniques available that are suitable for conventional computerized office applications. Some of these methods use the well-known Norman model [3]. Although this model lacks many aspects that might be appropriate for modeling business administration procedures, it was adopted as a start point in this research. Meanwhile, the development of the suggested system has followed ISO standard number TR18529 [7] which focuses on both different ergonomic aspects of human-system interaction and human-centered lifecycle process descriptions of interactive software systems.

Statistically, the study will try to investigate the above issues through giving specific tasks to the users, in addition to giving pre and post-experiment tests on the two systems for the purpose of measuring the users' attitudes so that their satisfaction, their ability for full recognizing the system, and the system performance verification could be noted. To avoid statistical errors caused through learn effects, permutations of the tasks and scenarios are taken into consideration. Finally, the study will highlight other usability aspects such as users' reactions, display screen design, terminology, consistency etc.

II. Designing the Experiment

A. Introduction

As mentioned, this contribution is aimed to statistically evaluate user-interfaces mainly used in business administration. Our intention through this evaluation is to reveal and compare the characteristics and abilities of a suggested prototype and a commercial software system. That is, designing an experiment that resembles programming the experiment in some ways. Each factor involved in the experiment can take a certain number of different values, and the experimental design employed specifies the levels of the one or many factors used in the experiment [8]. This evaluation is divided into two stages:

- In the first stage, experimental tasks such as establishing the questionnaire, collecting data, checking the stability of the questionnaire and variables by using (alpha) parameter and establishing scenarios, are carried out.
- In the second stage, it is concerned with acquiring the statistical raw data and focusing on the analysis of the raw data to highlight some usability differences between the investigated systems.

Results and guidelines achieved through this evaluation help and orient software system developers and user-interface designers in their tasks of both developing of user-oriented

Open Science Index, Industrial and Systems Engineering Vol:3, No:7, 2009 publications.waset.org/1414.pdf

Open Science Index, Industrial and Systems Engineering Vol:3, No:7, 2009 publications.waset.org/1414.pdf

business-administrative systems or optimizing existing ones, if possible. Since the commercial system serves as a reference, it has been possible to compare and analyze the statistical results accomplished using one of the statistical testing techniques available such as ANOVA and t-test. As it will be shown, the number of subjects was less than 30, and consequently the t-test technique is appropriate. Accordingly, our choice was for the simple t-test and not the complicated ANOVA, [9] and [10].

There are several factors that influence business administrative tasks such as time to learn and speed of performance etc. These factors do not only present standards for designing and developing interactive business-administrative systems, but they also serve as evaluation criteria. While some of these criteria might be subjectively obtained by means of questionnaires, others can be objectively achieved in an automated manner. For instance, in an objective evaluation, we measure users' motor-task performance like the time needed or the number of actions necessary to solve an offered scenario. Users' motor-task performance leans on discussing key-stroking or pointing times. In section IV "Comparative evaluation of the two systems", we are going to treat and highlight the major points of comparative usability testing for evaluating business-administrative user-interfaces.

B. Establishing the Questionnaire

The questions for subjective evaluation has been established after referring to different types of questionnaires: questionnaire of user interaction satisfaction (QUIS) [11], software usability measurement inventory (SUMI) [12], measuring usability of multimedia systems (MUMMS) [13], computer system usability questionnaire (CSUQ) [14], and questionnaires that deal with human-computer interaction [15]. Some of the above approaches such as QUIS avoids on-line procedures because they do not record specific comments about the system since they are often vital for usability testing. This problem has been overcome by monitoring the users' reaction to the system.

In order to carry out the tasks in our investigation, users receive trainings on both systems. Three well-trained and -educated scholars assisted the investigators in monitoring the subjects. The education and experience of the three assistants are as in the following: The first one has Master Degree in psychology with 20 years experience, the second is specialized in computer science with 8 years experience, and the third is specialized in Business Administration with 15 years experience.

C. Collecting Data

The raw data have been acquired through subjective and objective experimenting approaches. The subjective method is achieved by distributing a questionnaire to all users. Generally, this questionnaire covers the following aspects: administrative that consisted of 27 paragraphs following Likert scale, technical problems of investigated system, overall reactions to the system, screen, terminology and system information, learning, system capabilities, technical manuals and on-line help; and on-line tutorials. The other part of raw data has been obtained objectively and done by recording the users' interaction with the computer in an experimental session. It is worth mentioning that the users who use the core software package are 27 users. Their work is concentrated on processing all purchase transactions for a non-profit organization that deals with thousands of dollars annually. In our experiments, 6 users were trained to execute 3 scenarios on both systems where permutations were done among them.

D. Stability of Questionnaire and Variables

The questionnaire is tested by using (Alpha) parameter and found that, paragraphs depending on 5 degrees scale (Likart scale) got 85.89% and paragraphs depending on 7 degrees scale got 96.31%. The total average of the stability of the questionnaire for the two scales is 91.1% which satisfy the needs of the study. It is to note that questions within the general usability measure are highly related and specific questions also tend to be highly correlated with the overall.

In usability testing, a dependent variable is a factor determined by another variable called "the independent variable". In other words, the independent variable causes an apparent change in, or simply, affects the dependent variable. In analysis, researchers usually want to explain why the dependent variable has a given value. As mentioned, the values of a dependent variable in different settings are usually compared; whereas an independent variable is presumed to affect or determine a dependent variable. Usability engineers control/change the independent variable which causes the dependent variable to change as a result. In short, independent variables act as catalysts for dependent variables. That is, the independent variable is the "presumed cause" while dependent variable is the "presumed effect" of the independent variable [16]. In our investigation, the independent variables are the interfaces of the investigated (commercial) system and the prototype system; whereas the dependent represent variables the evaluation criteria: efficiency, effectiveness, and satisfaction, transparency, consistency, learnability, navigation etc.

E. Establishing Scenarios

Scenario-based methods both for statistical experiment design and analysis are not only beneficial for describing of people using technology in order to reshape their activities; but might be of great significance before a system is built and its impacts felt, [17] and [18]. During a usability testing session, the usability expert explains to the subject all operations related to the experiment in very simple and clear way. The following clarifies the scenarios used in these experiments:

1) The user is given a draft and is requested to process it all to the system. This scenario evaluates the process from the moment of switching on the PC.
2) The user is requested to do correction to the percent of the financial code after completing the first scenario immediately.
3) The user is requested to change the currency of the purchase.

It should be noted that the researchers have done permutations among the users to avoid statistical errors caused by learn effects.

Open Science Index, Industrial and Systems Engineering Vol:3, No:7, 2009 publications.waset.org/1414.pdf

F. Statistical Hypothesis

In order to discuss the results achieved by the usability testing, it is necessary to formalize and define statistical hypotheses for the evaluation. As mentioned, the dependent might be efficiency, effectiveness, satisfaction and error rate etc. Independent variables are both the sequences of the scenarios and the two business administration systems. For compensating the influence of such variables on the evaluation results, we have to permutate their sequences. The null and alternative hypotheses (H_0 and H_1) can be used to test whether there are significant differences between the mean values. The complete hypotheses' formulizations for the investigation are described in [19].

III. Decision-Making on Modifying the Existed Interactive Software System

Before making any decision on modifying the user-interfaces of an existed software system in an enterprise, it is significant to check the usability of that system. If the result achieved is unacceptable, it is valuable to enhance or even to replace the existed user-interfaces with newer, user-friendly ones. There are several usability criteria available as guidelines that are defined in international standards such as the ISO 9241-11 [20]. For a successful application of these guidelines, software system designers need to understand the design goals and benefits of each guideline, the conditions under which the guideline should be applied, the precise nature of the proposed solution, and any procedure that must be followed to apply the guideline. ISO 9241-11 consists of guidelines on usability, providing definitions of usability that is used in subsequent related ergonomic standards. Moreover, this ISO explains how to identify the information necessary to be taken into account when specifying or evaluating usability in terms of measures of user performance and satisfaction.

Guidance is given on how to describe the context of use of the product and the measures of usability in an explicit way. It does not only include an explanation of how the usability of a product can be specified and evaluated as part of a quality system, for example one that conforms to ISO 9001; but it also explains how measures of user performance and satisfaction can be used to measure how any component of a work system affects the quality of the whole work system in use as well.

Various evaluation criteria were used to test the usability of the business administration system such as effectiveness, efficiency, transparency, navigation, stress, confidence etc. Every subject has to answer various questions (see subsection II.B "Establishing the Questionnaire") leading to measure different subjective criteria such as:

- *Effectiveness:* It is accuracy and completeness with which specified users can achieve specified goals in particular environment [2]. Producing the result that is wanted or intended produces a successful result. The total average of effectiveness has a relatively low score; thus, it seems that some additional features should be added to activate the system. Also when the total evaluation is low, this means that the user needs more time than expected to carry out the given tasks.
- *Efficiency:* Efficiency is defined as the recourses expanded in relation to the accuracy and completeness of goals achieved. [2]; or the quality of doing something characterized with no waste of time or money. On this basis, evaluation is decided to improve the structure of the interfaces and to add a number of features before releasing the system to the users.
- *Satisfaction:* It is a tool to measure computer user's subjective satisfaction with the computer interface. It contains an overall measure of satisfaction and measures user satisfaction in four specific interface aspects: screen factors, terminology and system feedback, learning factors, and system capabilities [21].

Table I contains statistical data reflecting users' satisfaction with the original system regarding some business-administrative issues, and highlights the major criteria related to the original system that was measured from the users' point of views. It can be noted that the total percent of acceptance is 65.11% with a medium degree of acceptance. The total average of the mean is 3.25, which locates between undecided (3) and accepted (4), and more closely to undecided (3), which means that the users are not satisfied with the system.

While Table II and Table III include the statistical results of the subjective evaluation regarding users' satisfaction with the user-interfaces of the original system with different degree scales, Table IV includes some subjective data about the factor "Over all reactions to the system". Some very high scores in several questions are obtained, as some others are very low. Problems could be traced back to the structure of the interfaces. For examples, let us make some comments on Table II and Table IV. It is evident through the two tables that the lowest score obtained from the first scale (Table II) is 39.2% while the lowest one from the second scale (Table IV) is 46.43% and the grand total of the two scales is 54.02% which is classified in the weak phase. In other words, the user-interfaces of the original system (commercial system) require more modifications.

TABLE I
USERS' SATISFACTION WITH THE ORIGINAL SYSTEM REGARDING SOME BUSINESS-ADMINISTRATIVE ISSUES

Criteria	% Strongly Accept (5)	% Accept (4)	Undecided (3)	% Not accept (2)	% Strongly reject (1)	S. deviation	Mean	% of acceptance	Degree of acceptance
As a user, I feel it is easy to get the balances of the stocks	15.4	0	23.1	30.8	30.8	1.36	2.38	47.60%	Weak
There is no need to develop the system to have detailed reports	15.4	0	23.1	30.8	30.8	1.36	2.38	47.60%	Weak
The usage of the system has contributed to developing of the administrative system in the Logistics Department	19.2	46.2	11.5	19.2	3.8	1.14	3.58	71.60%	Strong

The system employment has contributed to developing of decision makings in the enterprise	15.4	34.6	34.6	15.4	0	0.95	3.50	70.00%	Strong
The business-administrative information delivered by the system is accurate and trustable	11.5	53.8	15.4	11.5	7.7	1.11	3.50	70.00%	Strong
The system employment leads to time saving	15.4	53.8	3.8	26.9	0	1.07	3.58	71.60%	Strong
The system employment leads to effort saving	15.4	53.8	3.8	26.9	0	1.07	3.58	71.60%	Strong
The system employment contributes to accuracy of reconciliations	11.5	46.2	19.2	23.1	0	0.99	3.46	69.20%	Medium
The system makes the appropriate reports available in the nick of time	15.4	30.8	19.2	26.9	7.7	1.23	3.19	63.80%	Medium
The system employment leads to information flow to high management and vice versa	11.5	46.2	30.8	11.5	0	0.86	3.58	71.60%	Strong
The system usage simplifies the operational procedures for obtaining goods from warehouse	11.5	30.8	19.2	30.8	7.7	1.20	3.08	61.60%	Medium
Total Average							3.25	65.11%	Medium

TABLE II
USERS' SATISFACTION WITH THE ORIGINAL SYSTEM REGARDING THE USER-INTERFACES USING A 5-DEGREE SCALE

Criteria	% Strongly Accept (5)	% Accept (4)	Undecided (4)	% Not accept (2)	% Strongly reject (1)	S. deviation	Mean	% of acceptance	Degree of acceptance
As a user of the system, the system is satisfied for me	3.8	7.7	15.4	26.9	46.2	1.14	1.96	39.2%	Weak
The system is so transparent	11.5	19.2	46.2	23.1	0	0.94	3.19	63.80%	Medium
The system is so accurate and assists in achieving the goals	11.5	38.5	11.5	34.6	3.8	1.17	3.19	63.80%	Medium
The other users are satisfied of the system	15.4	15.4	19.2	38.5	11.5	1.29	2.85	57.0%	Weak
The use of the system does not present a problem	7.7	15.4	30.8	34.6	11.5	1.12	2.73	54.6	Weak
Total Average							2.78	55.68	Weak

TABLE III
USERS' SATISFACTION WITH THE ORIGINAL SYSTEM REGARDING THE USER-INTERFACES USING A 7-DEGREE SCALE

Criteria	% +++ (7)	% ++ (6)	% + (5)	% 0 (4)	% - (3)	% -- (2)	% --- (1)	S. Deviation	Mean	Percent	Degree of Acceptance
The response time is accepted ---- not accepted	0	7.7	23.1	19.2	3.8	19.2	26.9	1.78	3.15	45.00	Very weak
The information displayed on the user-interface is suitable ---- unsuitable	7.7	26.9	15.4	15.4	11.5	7.7	15.4	1.96	4.19	59.83	Weak
Do you trust the system? Yes --- No	15.4	19.2	30.8	7.7	3.8	7.7	15.4	2.02	4.50	64.26	Medium
The acoustic feedback is meaningful ---- meaningless	0	7.7	26.9	26.9	3.8	7.7	26.9	1.77	3.42	48.84	Very weak

Open Science Index, Industrial and Systems Engineering Vol:3, No:7, 2009 publications.waset.org/1414.pdf

The user-interfaces do not need to modified agree ---- deny	11.5	11.5	3.8	7.7	23.1	15.4	26.9	2.10	3.27	46.69	Very weak
The easiness in using the user-interfaces is based on your experience agree ---- deny	38.5	19.2	19.2	3.8	3.8	7.7	7.7	1.97	5.31	75.82	Strong
It is easy to remember the sequences of operations agree ---- deny	15.4	3.8	23.1	3.8	19.2	15.4	19.2	2.09	3.69	52.69	Weak
When I use the system, I do not ask for a help agree --- deny	11.5	7.7	15.4	11.5	15.4	23.1	15.4	1.98	3.58	51.12	Weak
Total Average									3.89	55.53	Weak

TABLE IV
OVER ALL REACTIONS TO THE ORIGINAL SYSTEM

Criteria	% +++ (7)	% ++ (6)	% + (5)	% 0 (4)	% - (3)	% -- (2)	% --- (1)	S. Deviation	Mean	Percent	Degree of Acceptance
Wonderful--------------Terrible	19.2	15.4	38.5	3.8	11.5	7.7	3.8	1.68	4.88	69.69	Medium
Satisfying ---------- Frustrating	7.7	19.2	34.6	0	0	30.8	7.7	1.97	4.12	58.83	Weak
Stimulating------------------ Dull	3.8	3.8	34.6	7.7	3.8	15.4	30.8	1.97	3.27	46.43	Weak
Easy----------------------Difficult	7.7	11.5	26.9	3.8	7.7	19.2	23.1	2.08	3.58	51.12	Weak
Flexible----------------------Rigid	7.7	3.8	26.9	7.7	3.8	26.9	23.1	2.00	3.31	47.27	Weak
Total Degree										54.67	Weak
Grand Total										54.02	Weak

IV. COMPARATIVE EVALUATION OF THE TWO SYSTEMS

After having developed and implemented the suggested system (prototype), we had to verify through experimental tests whether the developed prototype fulfilled the user requirements on the systems user-interfaces or not. This can be achieved through a comparison between the two systems by collecting data about their performance by using SPSS [23], and then draw conclusions from statistical analyses. The manner in which sample data are collected called an experimental design which is crucial to an investigation.

As previously mentioned, the comparative usability testing reveals how much are the desired requirements on a new prototype are fulfilled. Usability testing builds an important ingredient of the iterative design process, which consists of these core stages [6]:

- Studying the users and their tasks as part of task analysis.
- Making a prototype early in the design phase and then reviewing it with expert users.
- Testing the prototype usability with typical users.
- Correcting the prototype, testing it again and so until the desired result is achieved (see Fig. 2).

Unlike to subjective evaluation techniques accomplished through questionnaires and surveys, objective evaluation techniques that are exclusively applied in this comparative usability testing, are achieved in an automated manner. (see subsection II.C "Collecting Data"). For instance, objective evaluation measures users' motor-task performance like the time needed or the number of actions necessary to solve an offered scenario. Users' motor-task performance leans on predicting key-stroking or pointing times. The following are the criteria [22] that have been adapted to the needs of this investigation which enable us to compare our two systems objectively: Time to complete a task, percent of time completed, percent of task completed per unit time, ratio of success to failures, time spent in errors, percent of number of errors, percent of number of competitors better than it, number of commands used, frequency of help and documentation, percent of favorable/unfavorable user comments, number of repletion's of failed commands, number of runs of successes and of failures, number of times interface misleads the user, number of good and bad features recalled by users, number of available commands not invoked, number of regressive behaviors, number of users preferring your system, number of

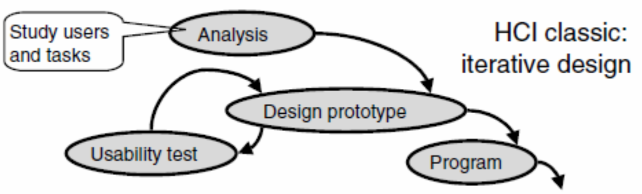

Fig. 2 Iterative design process [6]

Open Science Index, Industrial and Systems Engineering Vol:3, No:7, 2009 publications.waset.org/1414.pdf

times users need to work around a problem, number of times the user is disrupted from a work task, number of times user loses control of the system, and number of times user expresses frustration or satisfaction. As previously stated, t-test is employed in this phase of study as the number of population is less than 30.

It is tangible that computers do more than just providing information and offering services to people to use [24]. Caroll [25] notes that the design of computing systems is part of an ongoing cycle in which new technologies raise new opportunities for human activities; as people's tasks change in response to these opportunities. Scenario-based methods both for designing and analyzing are not only beneficial for describing of people using technology in order to reshape their activities; but might be of great significance before a system is built and its impacts observed, [26] and [18]. Here, the subjects should carry out various business-administrative tasks organized in three scenarios. As the case with the independent and dependent variables, the scenarios were permutated in order to avoid undesirable statistical effects leading to falsification of the experiments results. The three scenarios are:

- *First Scenario:* The user is given a draft of the transaction and he is requested to process it to both systems.
- *Second Scenario:* The user is requested to modify the percent of the financial codes on both systems.
- *Third Scenario:* The user is requested to change the type of the currency of purchase on both systems. It can be stated that the global score has an average value of 50 in a normal distribution.

Table V, Table VI and Table VII indicate comparisons between the two systems by pointing out of statistical results for the different criteria used in the three scenarios [19]. In order to illuminate the statistical meaning of the tables, consider the criterion "No. of unfavorable user comments" in Table VII for the third scenario. For the commercial system, Mean is equal 1.00, T equals 3.873 and its significance is 0.012 which is less than 0.05, locating in the rejection region. In other words, the users express unfavorable comments regarding the commercial system. According to the suggested prototype, Mean equals 0.5, T is equal to 2.236 and the significance is 0.076 which greater than 0.05. As it is located in the acceptance region, this means that the users express positive comments about the prototype. It is obvious that there are interrelations between the confidence intervals for some phrases when comparing the two systems. This reflects the importance of the phrase from the point view of the user. Additionally, T is not calculated for some phrases because the standard deviation is equal to zero, that means the T value is not identified and the value is located on the population mean.

TABLE V
STATISTICAL DATA ACQUIRED BY THE COMPARATIVE USABILITY TEST (1ST SCENARIO)

Criteria	Program X						Prototype					
	Mean	SD	T	Sig	Confidence Interval		Mean	SD	T	Sig	Confidence Interval	
					Lower	Upper					Lower	Upper
Times of use of mouse	44.00	7.950	13.557	.000	35.66	52.34	19.83	2.137	22.734	.000	17.59	22.08
Time to execute the task/minutes	11.17	2.483	11.015	.000	8.56	13.77	4.83	.753	15.727	.000	4.04	5.62
No. of commands used	30.00	.00(a)					17.00	.000(a)				
No. of favorable user comments	.83	1.169	1.746	.141	-.39	2.06	1.00	.000(a)				
No. of unfavorable user comments	1.33	1.211	2.697	.043	.06	2.60	.67	.816	2.000	.102	-1.9	1.52
No. of repetitions of failed commands.	.83	.983	2.076	.093	-.20	1.87	.50	.548	2.236	.076	-.07	1.07
No. of runs of successes	1.33	.816	4.000	.010	.48	2.19	1.00	.000(a)				
No. of runs of failures	.50	.548	2.236	.076	-.07	1.07	.17	.408	1.000	.363	-.26	.60
No. of times interface misleads the user.	1.33	.816	4.000	.010	.48	2.19	.17	.408	1.000	.363	-.26	.60
No. of available commands not invoked	.33	.816	1.000	.363	-.52	1.19	.00	.000(a)				
No. of regressive behavior	1.33	.816	4.000	.010	.48	2.19	.50	.548	2.236	.076	-.07	1.07
No. of times the user is disrupted from a work task	1.50	1.975	1.861	.122	-.57	3.57	.17	.408	1.000	.363	-.26	.60
No. of times the user loses control of the system	.83	.408	5.000	.004	.40	1.26	.50	.548	2.236	.076	-.07	1.07

Open Science Index, Industrial and Systems Engineering Vol:3, No:7, 2009 publications.waset.org/1414.pdf

No. of times user expresses frustration	1.50	.837	4.392	.007	.62	2.38	.50	.548	2.236	.076	-.07	1.07
No. of times the user expresses satisfaction	1.17	.408	7.000	.001	.74	1.60	.83	.408	5.000	.004	.40	1.26

TABLE VI
STATISTICAL DATA ACQUIRED BY THE COMPARATIVE USABILITY TEST (2ND SCENARIO)

Criteria	Program X						Prototype					
	Mean	SD	T	Sig	Confidence Interval		Mean	SD	T	Sig	Confidence Interval	
					Lower	Upper					Lower	Upper
Times of use of mouse	26.67	7.448	8.771	.000	18.85	34.48	9.17	1.835	12.237	.000	7.24	11.09
Time to execute the task/minutes	4.67	1.862	6.139	.002	2.71	6.62	2.67	.816	8.000	.000	1.81	3.52
No. of commands used	19.00	.000(a)					5.00	.000(a)				
No. of favorable user comments	1.00	.894	2.739	.041	.06	1.94	.17	.408	1.000	.363	-.26	.60
No. of unfavorable user comments	1.67	.816	5.000	.004	.81	2.52	.33	.516	1.581	.175	-.21	.88
No. of repetitions of failed commands.	.50	.548	2.236	.076	-.07	1.07	.17	.408	1.000	.363	-.26	.60
No. of runs of successes	1.00	.000(a)					1.00	.000(a)				
No. of runs of failures	.83	.408	5.000	.004	.40	1.26	.17	.408	1.000	.363	-.26	.60
No. of times interface misleads the user.	1.33	.516	6.325	.001	.79	1.88	.00	.000(a)				
No. of available commands not invoked	.33	.516	1.581	.175	-.21	.88	.17	.408	1.000	.363	-.26	.60
No. of regressive behavior	1.33	.516	6.325	.001	.79	1.88	.50	.837	1.464	.203	-.38	1.38
No. of times the user is disrupted from a work task	.67	.816	2.000	.102	-.19	1.52	.17	.408	1.000	.363	-.26	.60
No. of times the user loses control of the system	1.17	.753	3.796	.013	.38	1.96	.17	.408	1.000	.363	-.26	.60
No. of times user expresses frustration	1.33	.816	4.000	.010	.48	2.19	.33	.816	1.000	.363	-.52	1.19
No. of times the user expresses satisfaction	1.00	.000(a)	8.771	.000			1.00	.000(a)	12.237	.000		

TABLE VII
STATISTICAL DATA ACQUIRED BY THE COMPARATIVE USABILITY TEST (3RD SCENARIO)

Criteria	Program X						Prototype					
	Mean	SD	T	Sig	Confidence Interval		Mean	SD	T	Sig	Confidence Interval	
					Lower	Upper					Lower	Upper
Times of use of mouse	29.17	5.115	13.967	.000	23.80	34.53	9.33	1.966	11.626	.000	7.27	11.4
Time to execute the task/minutes	6.50	1.871	8.510	.000	4.54	8.46	1.83	.408	11.000	.000	1.40	2.26
No. of commands used	21.00	.000(a)					5.00	.000(a)				
No. of favorable user comments	.83	.753	2.712	.042	.04	1.62	1.00	.632	3.873	.012	.34	1.66
No. of unfavorable user comments	1.00	.632	3.873	.012	.34	1.66	.50	.548	2.236	.076	-.07	1.07

Open Science Index, Industrial and Systems Engineering Vol:3, No:7, 2009 publications.waset.org/1414.pdf

No. of repetitions of failed commands.	.67	.516	3.162	.025	.12	1.21	.17	.408	1.000	.363	-.26	.60
No. of runs of successes	1.00	.000(a)					1.00	.000(a)				
No. of runs of failures	.67	.516	3.162	.025	.12	1.21	.50	.837	1.464	.203	-.38	1.38
No. of times interface misleads the user.	1.17	.753	3.796	.013	.38	1.96	.17	.408	1.000	.363	-.26	.60
No. of available commands not invoked	.33	.516	1.581	.175	-.21	.88	.00	.000(a)				
No. of regressive behavior	1.17	.408	7.000	.001	.74	1.60	.50	.548	2.236	.076	-.07	1.07
No. of times the user is disrupted from a work task	.83	.983	2.076	.093	-.20	1.87	.17	.408	1.000	.363	-.26	.60
No. of times the user loses control of the system	1.17	.408	7.000	.001	.74	1.60	.00	.000(a)				
No. of times user expresses frustration	1.17	.408	7.000	.001	.74	1.60	.67	.516	3.162	.025	.12	1.21
No. of times the user expresses satisfaction	1.17	.408	7.000	.001	.74	1.60	.50	.548	2.236	.076	-.07	1.07

Open Science Index, Industrial and Systems Engineering Vol:3, No:7, 2009 publications.waset.org/1414.pdf

V. Conclusion

It has been shown how we can evaluate user-interfaces of different business-administrative systems through comparing them with each other. According to Shneiderman et al. [21], there are several factors that influence the business-administrative tasks carried out through computerized systems accessed by graphic user-interfaces such as time to learn, speed of performance, extending global reach, maximizing impact and integration, responding to demand, retention over time, rate of errors, and subjective satisfaction. This study was aimed at investigating the effect of using computerized business administration on the achievement of users (see subsections II.B and II.C) dealing with various business-administrative tasks such as purchase, transactions etc.

The raw data of the experiments were handled statistically by using the Student's test (t-test) and then analyzed by SPSS. Through the usability test, we have found the suggested prototype system is superior to the commercial one because various modern computer graphics and multimedia issues were taken into consideration to cover the drawbacks of the commercial system. The two systems were compared with each other by means of an experimental testing, in which some selected usability criteria such as satisfaction, effectiveness, efficiency etc. were used. The investigation has shown that the employment of user-centered user-interfaces in business administration has significant effects on staff achievements in an enterprise. For further investigation of the usability testing, it is recommended that we carry out more experimental studies on the role of the user-centered computer systems in business administration.

It is necessary to note that, before the employment of any product, the organization should take into consideration the usability and applicability of this product; statistical methods can be good indicator for the purpose of evaluating by reasoning and guiding the purposeful collection and analysis of data towards the continuous improvement of any process. Results and guidelines achieved through this evaluation help and orient software system developers and user-interface designers in their tasks of both developing of newer computerized systems for business administration with user-centered user-interfaces or optimizing the existing ones.

References

[1] I. Sommerville, *Software Engineering* (8th Edition). Addison Wesley, pp. 20-42, 2007.

[2] E. A. Hartmann and A. Westerwick, “Shopfloor Systems Based on Human Skill and Experience,” in *Proc. 7th IFAC/IFIP/IFORS/IEA Symposium on Analysis; Design and Evaluation of Man-Machine Systems*, Kyoto, Japan, Sept. 16-18, 1998.

[3] A. Dix, J. Finlay, G. Abowd, R. Beale, *Human Computer Interaction*, 2nd Edition.Prentice Hall – Europe, 1998.

[4] H. Yajima, Y. Yoshikawa, M. Tamano, R. Itsuki and T. Hirose, “*Information Integration of Heterogeneous Multimedia Information for Strategic Planning System,*” in Proc. of the 7th International Federation of Automatic Control – Man Machine Systems (IFAC-MMS). KYOTO, JAPAN, September, 16-18, 1998.

[5] R. G. Bias and D. J. Mayhew, *Cost-Justifying Usability: An Update for the Internet Age*. Morgan Kaufmann, Elsevier, pp. 19, 2005.

[6] S. Lauesen, *User Interface Design: A software Engineering Perspective.* Harlow: England: Pearson-Addison-Wesley, 2005, pp. 41-66.

[7] ISO 13407: Human-centered design processes for interactive systems [Online]. Available: http://www.usabilitynet.org/tools/13407stds

[8] Y. Dodge, *The Concise Encyclopedia of Statistics*. Springer, 2008.

[9] P. L. Gardner, "Discusses assumptions of the t-test," *Scales and Statistics: Review of Educational Research*, vol. 45, pp. 43-57, 1975.

[10] R. Johanson and G. Bhattacharyhya, *Statistics Principles and Methods*. 3rd Edition, John Wiley & Sons, Inc, 1996.

[11] J. P. Chin, V. A. Diehl and K. L. Norman, K. L. „*Development of an instrument measuring user satisfaction of the human-computer interface*,” in proc. of CHI ’88 human factors in computing systems, pp. 213–218, New York: ACM Press, 1988.

[12] SUMI: Software Usability Measurement Inventory [Online]. Available: http://sumi.ucc.ie/whatis.html

[13] MUMMS: Measuring the Usability of Multi-Media Software [Online]. Available: http://www.ucc.ie/hfrg/questionnaires/mumms/index.html

[14] 12 J. R. Lewis, “IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use,” *International Journal of Human-Computer Interaction*, vol. 7, no. 1, pp. 57-78, 1995

[15] A. A. Ozok, “Survey design and implementation in HCI,” *in handbook The Human–Computer Interaction Handbook Fundamentals, Evolving Technologies, and Emerging Applications*, 2nd ed., A. Sears and J. Jacko, Ed. New Yoork: Taylor & Francis Group, 2008, pp. 1151-1169.

[16] N. A. Streitz, "*Cognitive compatibility as a central issue in human-computer interaction: Theoretical framework and empirical findings*", in G. Salvendy (Ed.), Cognitive engineering in the design of human-computer interaction and expert systems. Amsterdam: Elsevier, 1987, pp. 75-82.
[17] M. B. Rosson, S. Maass, W. A. Kellog, "*The designer as user: Building requirements for design tools from design practice,*" Communications of the ACM, vol. 31, no. 11, pp. 1288–97, 1989.
[18] K. Weidehhaupt, K. Pohl, M. Jarke and P. Haumer, "Scenarios in system development," *Current practice. IEEE Software*, vol. 15, no. 2, pp. 34–45, 1998.
[19] I. Adwan, "*Applying and developing usability engineering for the administration of Procurement and Warehousing Processes in "UNRWA" by Using the Software package - (Reality)*," MBA thesis, Institute Business and Economic Studies, Al-Quds Univ., Abu Dies, Jerusalem, 2006.
[20] ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability, 1998.
[21] B. Shneiderman, C. Plaisant, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. 4th Edition, Addison Wesley Longman, 2004.
[22] K. Bohmann. (2000, July 4). User Performance Metrics [Online]. Available: http://www.bohmann.dk/articles/user_performance_metrics.html
[23] J. Pallant, *SPSS Survival Manual: A Step By Step Guide to Data Analysis Using SPSS for Windows*, Open University Press, 2007.
[24] M. B. Rosson and J. M. Carroll, *Usability Engineering: Scenario-Based Development of Human-Computer Interaction*, Morgan Kaufmann Publishers, 2002.
[25] J. M. Caroll, *Making Use: Scenario-Based Design of Human-Computer Interactions*, Cambridge, MA: MIT Press, 2000.
[26] M. B. Rosson, S. Maass and W. A. Kellog, "*The designer as user: Building requirements for design tools from design practice*," Communications of the ACM, vol. 31, no. 11, pp. 1288–97, 1989.

Salaheddin Odeh received his masters degree in electrical engineering, specialization area computer and control engineering, from the University of Stuttgart and the PhD degree from the University of Kassel; both universities are in Germany. He is currently an assistant professor in the Department of Computer Engineering, Faculty of Engineering, Al-Quds University in Palestine. In 1999, for his doctoral thesis, he was the recipient of the first prize of the Association of German Engineers (VDI) of Northern Hesse in Germany for the best technical-scientific research.
Ibrahim Othman Adwan received his master degree in Business Administration from Al-Quds University – Palestine. He is currently filling the post of Senior Personnel Assistant (Entitlements) at UNRWA – Administration Department. His master thesis is about *"Applying and developing usability engineering for the administration of Procurement and Warehousing Processes in "UNRWA" by Using the Software package - (Reality)*".

Open Science Index, Industrial and Systems Engineering Vol:3, No:7, 2009 publications.waset.org/1414.pdf