Computers in Human Behavior 61 (2016) 463–471

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

Exploring user satisfaction for e-learning systems via usage-based metrics and system usability scale analysis

Nouzha Harrati [a,b,*], Imed Bouchrika [b], Abdelkamel Tari [a], Ammar Ladjailia [b]

[a] Department of Computer Science, University of Bejaia, 6000, Algeria
[b] Faculty of Technology and Science, University of Souk Ahras, 41000, Algeria

ARTICLE INFO

Article history:
Received 25 December 2015
Received in revised form
24 February 2016
Accepted 15 March 2016

Keywords:
e-Learning
Moodle
Usability
Usability evaluation

ABSTRACT

The use of e-learning technology is incontestably recognized as an important and integral part of the educational process. Considerable research studies are carried out in order to apprehend how effective and usable e-learning systems. In this paper, an empirical-based study is conducted to explore how lecturers interact with an e-learning environment based on a predefined task model describing low-level interactions. Client-side log data is collected from university lecturers from the Electrical and Computer Science departments. Subsequently, data analysis is conducted to infer the usability degree from the estimated usage metrics together with further exploratory analysis from user feedback via System Usability Scale. Experimental results reveal that the System Usability Scale score is not a sufficient measure to express the true acceptance and satisfaction level of lecturers for using the e-learning systems. The evaluation must be fulfilled in tandem with analyzing the usage metrics derived from interaction traces in a non-intrusive fashion. The proposed approach is a milestone towards usability evaluation to improve the acceptance and user experience for academic staff and students.

1. Introduction

The unprecedented growth of internet-based technology has led to the advent of numerous approaches devoted to the area of education manifested in the use of e-learning systems. The pace of e-learning is getting remarkably higher as most of the educational institutions have already installed web-based systems for offering online courses (Garrison, 2011; Guri-Rosenblit, 2006). These often complement traditional methods enabling students to engage from any place with their learning through various materials alongside or instead face-to-face teaching delivery (Guri-Rosenblit, 2006). The European commission defines the e-learning process as the use of Internet and multimedia technologies to improve the quality of teaching through providing access to resources and educational services as well as enabling remote evaluation, exchange and collaboration between students and lecturers (Dominici & Palumbo, 2013; Nichols, 2008). The main motives behind the use of e-learning technology are to enforce a profound change in the way teaching takes place in educational institutions from the still widespread transmissive approach towards the more participated, self-regulated and interactive methods that are considered as important for improving the learning outcomes (Garrison, 2011; Persico, Manca, & Pozzi, 2014). Interestingly, e-learning platforms are being recently introduced in schools, public administrations and corporations to increase the learning opportunities and overcome the drawbacks of traditional teaching (Stoffregen, Pawlowski, & Pirkkalainen, 2015; Violante & Vezzetti, 2015).

As more and more universities worldwide have opted to use e-learning environments for their course delivery, research in e-learning systems have attracted considerable interest in order to apprehend how effective and usable e-learning systems in terms of principles related to human computer interaction (Bringula, 2013; Escobar-Rodriguez & Monge-Lozano, 2012; Navimipour & Zareie, 2015). Positive user experience is of prime importance for educational learning systems playing vital role for the acceptance, satisfaction and efficiency of academic institutions. This is mainly because the availability of technological infrastructures and systems is not adequate to enforce the uptake of new educational approaches from both sides of the lecturer or the learner (Laurillard, Oliver, Wasson, & Hoppe, 2009; Persico et al., 2014; Phillips, McNaught, & Kennedy, 2012). The usability nature of e-learning software products is a key characteristic to achieve the

* Corresponding author.
E-mail addresses: n.harrati@univ-soukahras.dz (N. Harrati), imed@imed.ws (I. Bouchrika).

http://dx.doi.org/10.1016/j.chb.2016.03.051
0747-5632/

acceptance of academic users regardless of their background, experience or orientation. Usability is defined as the extent to which a product can be easily used by specified users to achieve certain goals with effectiveness, efficiency and satisfaction. The satisfaction part is related to how the users believe or feel positively that the system meet their requirements (Capece & Campisi, 2013; Lee, Kim, & Lee, 1995; Yeh & Lin, 2015). Meanwhile, other researchers have defined satisfaction as the gap between the expected gain and the actual again when using the system (Tsai, Yen, Huang, & Huang, 2007). Software systems are valued on the basis of its graphical interface and the related power of communication and expression for the implemented functionalities (Cassino, Tucci, Vitiello, & Francese, 2015). It is no doubt that usability is now recognized as an important software quality attribute, earning its place among more traditional attributes such as performance, robustness and security.

There is a dearth of studies and approaches devoted for the exploratory evaluation of the acceptance and usability aspect by university lecturers for using e-learning applications. Motivated by the fact that the process for introducing e-learning systems is bound to have a slow and complex trend (Persico et al., 2014) which needs to be understood and evaluated beyond the use of just summative ways, hence we explore in this paper an empirical-based study to assess the satisfaction level of how lecturers interact with an e-learning environment system based on a pre-defined task model describing low-level interactivity details. The main thrust of this research is to evaluate the usability of the e-learning platform as usability is considered a vital attribute for the adoption of educational systems by lecturers. An online automated system for formalizing user interaction with a given system guided through a set of rules describing certain goals to be achieved by the end user is setup for usability practitioners. The task model is mainly utilized to capture all the interactions and navigation path to be carried out by the university staff. Empirical client-side log data is collected from university lecturers from the Electrical and Computer Science departments participating within the usability evaluation of the e-learning system in a non-intrusive fashion without the need to install additional tools. The Moodle e-learning platform is used as the case study for this research. Subsequently, data analysis is conducted to infer the usability level. This is carried out in compliance with the defined task model and usability metrics describing efficiency of use. Regardless of the fact that users have expressed higher satisfaction scores through the System Usability Scale (SUS) (Brooke, 1996), empirical results performed to inspect the usability of the e-learning platform have revealed that potential reasons to impede the adoption of new technologies within the teaching process is primarily related to the complex nature of the software interface where the majority of lecturers failed to complete simple tasks.

This paper is organized as follows. The next section outlines the existing approaches and studies related to the evaluation of user satisfaction for using e-learning systems. The theoretical description of the presented approach for quantifying the usability is described in Section 3. Section 4 is devoted to show the experimental results attained for the evaluation of the Moodle Platform. Finally, discussion and conclusions are drawn.

2. Literature review

Evaluation of e-learning applications in terms of user experience and satisfaction has received recently considerable attention from the research community in order to assess and quantify the satisfaction and effectiveness level for both students and lecturers. This is due to the increasing concern that despite the wide use of e-learning technologies, the intended impact on education is not achieved (Asarbakhsh & Sandars, 2013; Phillips et al., 2012). Ivory et al. (Ivory & Hearst, 2001) argued that the automation of usability evaluation for software systems would help to increase the coverage of testing as well as reduce significantly the costs and time for the evaluation process. Hornbæck (Hornbæk, 2006) reviewed an extensive list of studies related to usability evaluation covering over 180 published papers examining critically the practice of measuring usability of software systems. The main aim of the conducted study was to understand and explore the challenges facing the reliability as well as the validity of usability evaluation methods. Fernandez (Fernandez, Insfran, & Abrahão, 2011) surveyed the recent studies related to usability evaluation where they have categorized the different methods into broadly two main classes; empirical and inspection methods. The study reviewed by Fernandez was primarily aimed to investigate the applicability of the usability evaluation techniques for web applications.

2.1. E-learning evaluation models

There are a number of methods and theories in the literature for understanding, predicting, and assessing the interaction process with its involved parts including personal factors, behavior, and the environment. In order to assess the user acceptance for technological products, one of the most well established models is the Technology Acceptance Model (TAM), which was proposed by Davis (1989). The TAM is tailored to include questions to explore two aspects of the user satisfaction which are: perceived ease-of-use and perceived usefulness. The ease of use refers to how users believe that adopting a particular technological product would require no effort and hassle to use it (Davis, Bagozzi, & Warshaw, 1989). The perceived usefulness concerns the degree to which a user believes that using a particular software system would improve their job performance. The Technology Acceptance Model has been used in various studies to assess the factors affecting individual's use of technology (Venkatesh & Davis, 2000). For research studies related to assess the usability aspect of the Moodle e-learning platform, Persico (Persico et al., 2014) employed the Technology Acceptance Model to investigate the willingness of university users for the adoption of e-learning systems. Evaluation is based on three dimensions including usefulness, ease of use and effectiveness. Escobar-Rodriguez (Escobar-Rodriguez & Monge-Lozano, 2012) analyzed how university students use the Moodle platform in order to determine and understand the factors which might influence their intention to use the platform. The Technology Acceptance Model is used to assess the usability of the system in terms of perceived usefulness and ease of use against actual usage behavior.

There are other related models and theories such as The System Usability Scale (SUS) which was proposed mainly for the evaluation of web applications for two aspects; the learnability and usability. The SUS is a well-researched and widely used questionnaire for assessing the usability of mostly web applications. Surprisingly, only a few studies in the literature have used SUS to evaluate the perceived usability of e-learning management systems (Orfanou, Tselios, & Katsanos, 2015). The first study of using the SUS for e-learning systems was conducted by (Renaut, Batier, Flory, & Heyde, 2006) to inspect usability problems for the SPIRAL platform. The researchers employed the SUS scale as a post-assessment of the usability of the software system reporting an average score of 72% for the participating university lecturers who described the platform as positively easy to use. In (Simões & de Moraes, 2012), Simoes examined the usability of the Moodle e-learning platform using three different evaluation methods including the SUS questionnaire to assess user's satisfaction for a sample size of 59 students. The authors concluded that the SUS is an effective tool for exploring the usability aspect without reporting the obtained SUS score. Marco (Marco, Penichet, & Gallud, 2013) proposed a way of remote collaboration in real time within the platform Moodle through the use of Drag & Share. The collaborative tool enables sharing and synchronization of files. The efficiency of users was quantified using the time taken for task completion meanwhile user satisfaction was assessed using the SUS questionnaire with a reported score of 89.5%.

2.2. Factors affecting e-learning

Li, Duan, Fu, and Alford (2012) argued that among the factors that directly influence the intention of users towards using an e-learning system include service and course quality, perceived usefulness, perceived ease of use in addition to self-efficacy. The system functionality and response have indirect influence on the user satisfaction meanwhile the study reported that the system interactivity affects insignificantly the users' acceptance. Bringula (2013) had used a hierarchical regression analysis to determine the usability factors that have an influence on the interactivity for a university e-learning web portal. In their investigation, it was reported that age, commitment to use the web portal in addition to the information content are found to influence the usability of the e-learning system. Meanwhile, the commitment of university users is considered as a positive thrust to entice older people to get engaged with the e-learning portal. In a study conducted recently by Orfanou et al. (2015) involving 769 university students to assess their satisfaction for using the platform Moodle, an average SUS score of 76.27% is reported. Various SUS attributes were investigated in the context of evaluating the e-learning platform. Factors include: gender, age, prior experience, Internet self-efficacy, attitude towards the Internet in addition to usage frequency.

2.3. Automated usability evaluation

Analytical approaches involve usability practitioners to manually examine a graphical user interface in order to detect usability deficiencies via inspecting usage test cases or analyzing the results of questionnaires. Although such methods are known to be laborious and very expensive, they often yield results that are biased by the acquisition environment or experts' subjectivity. Alternatively, several automated evaluation methods are conceived for auto discovery of usability faults at the same time avoiding the drawbacks in terms of reducing costs and time through liberating usability experts from conducting manually repetitive tasks in addition to increase the coverage of tested features (Quade, Lehmann, Engelbrecht, Roscher, & Albayrak, 2013). Furthermore, because of the immense volume of data acquired from usability evaluation, the total or partial use of automated methods can be very beneficial during the development of web applications (Cassino et al., 2015; de Santana & Baranauskas, 2015). However, the majority of the surveyed research studies are purely based on manual or statistical analysis of recorded activity data for the participants. Methods for usability evaluation are conventionally grouped into two main categories; the first class is based on analyzing the graphical interface through reading the source code of the website to examine the content and structure of the application. Cassino et al. (Cassino & Tucci, 2011) assessed the source code to infer the design model of the interface and the interaction styles implemented on every page of the website to generate a quantitative report of the evaluation based on heuristic factors. Meanwhile, other methods rely on examining the usage data i.e. logs. The user logs used for usability evaluation are captured at either the server-side or the client-side. Many studies advocate that logging techniques are proven to be more reliable and efficient in terms of providing useful usability insights for the evaluators (de Santana & Baranauskas, 2015).

3. Methodology

3.1. Context of the study

The work described in this paper is carried out to explore the factors related to the cause of university lecturers abstaining from adopting the use of e-learning systems. From a theoretical point of view, the main objective is to determine whether perceived user satisfaction and acceptance are entirely dependent on or related to the quantified usability of technological products. From the experimental side, the evaluation process is conducted on the Moodle platform which is an open source e-learning management system being used worldwide by a large number of academic institutions including the universities where the study is conducted. The stable version 2.9.2 of Moodle is installed on a remote accessible web server with the usage logger scriplet integrated within the HTML pages of the website from the administration section. Upon testing the application, participants are not required to install any software apart from using their preferred browser to test the interface. All actions and events performed by the users are recorded automatically and non-intrusively into the log database. For legal and privacy concerns, participants have expressed their written consent in advance that their usage logs are recorded anonymously for improving user experience and analyzing usability of the given system. The usability logs dataset is made publicly available at the address: www.usability.ws/dataset.

3.2. Participants

For the experimental setup, an evaluation session is organized with a list of 50 lecturers from the Computer Science and Electrical Engineering departments across different universities. The age distribution of the users is ranging between 26 and 65 years with an average of 37 years old as illustrated in Table 1. There are 27 male participants and 23 females. 30% of the users are invited to conduct the experiment within a laboratory meanwhile the rest of lecturers are being asked to use the e-learning platform remotely from their workplace through emailing them a special link. For the educational background of the participants, 14% are full university professors meanwhile 36% are lecturers who completed their doctorate degrees. The rest of users are assistant lecturers who are still working towards the completion of their PhDs.

3.3. Experience-based metrics estimation

In order to quantify the metrics related to inspecting the usability aspect for the e-learning Moodle platform, the proposed system consists of three main modules: i) Task Modeling ii) Usage Tracking, iii) Metrics Estimation. The implementation of the proposed system is made available online at the address: http://www.usability.ws. An overview describing the developed framework is presented in Fig. 1 showing the three different components used in this research.

3.3.1. Task-based descriptor

A tree-based graphical representation is proposed for constructing a task model that should describe the tasks, actions and goals to be achieved by the participant. The resulting task model tree represents all interactions that a user can perform on a given web interface. Using the tree-based visual notation, the task model is an ordered hierarchy of tasks or other elements to be performed in order to satisfy a specific goal for a task. In order to enable

Table 1
Gender & age distributions of participants.

	Age distribution				Gender		Academic qualification		
	25–29	30–39	40–49	50–65	Male	Female	Professor	Lecturer	Assistant Lect.
Number	16	17	9	8	27	23	7	18	25
Percentage %	32	34	18	16	54	46	14	36	50

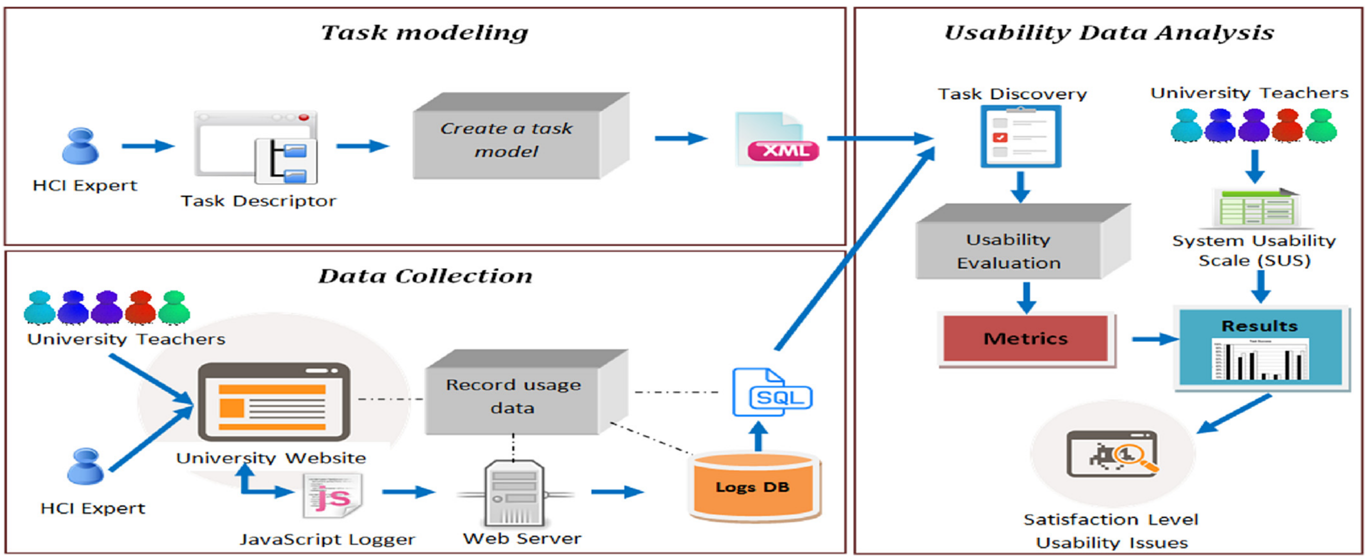

Fig. 1. Proposed framework for remote usability evaluation.

automation at later stages, goals for actions should have a way to infer automatically whether a task is completed successfully based on conditions and events triggered either by the user or the system. A task which is described only by name, consists of actions to be performed to achieve a specific goal. This can be a basic task consisting mainly of simple actions such as clicking a submit button, page scrolling or typing a text into a text field. For each basic action, there should be a mapping to an event caused by performing an action. In addition, it can be a complex task composed from other subtasks and advanced control blocks such as filling a checkout payment form for an online shopping cart containing many widgets with a number of options and conditions to be verified. Further, tasks can be combined to describe higher level tasks and processes. For simplicity of modeling, only tasks concerning the user interactivity with the system are considered inversely to system tasks executed either by the browser or on the server side (Guerrero Garcia, Vanderdonckt, & Gonzalez Calleros, 2008) such as checking or processing credit card payments. Various control blocks are employed for expressing the temporal relationship for task children which determines the number and logical order in which the subtasks must be performed by the user to achieve a particular goal. Control blocks include sequence, iteration and choice.

For the context of this research study, Fig. 2 shows the task model constructed for using the Moodle e-learning platform. The task model contains four consecutive high level tasks which describe the login activity and the process of uploading a teaching resource by university lecturers within an assigned course for their students. During the first task, the user is presented with the front page containing a login link at the top header which would take the user to the login page with a form asking the user for a username and password. This task is marked as successful when the user is logged in and redirected to the *my account* page. Subsequently, participants would start the second task which is browsing to the course and enabling the editing mode. Thirdly, the user is required to click on the *add resource* option, choose file and click the submit button. In the last task, users are asked to type a title for the file and upload the given *pdf* file from their local computers. Afterwards, they click the *"save and return to course"* button to complete the experiment session.

3.3.2. Data collection

For usability analysis, it is typical to automatically record clicks, page views and duration in order to determine conversion rates and website traffic. For the course of this research, a JavaScript program is implemented to log all user activities performed when browsing the e-learning platform to get its usability inspected. The proposed tool is integrated by appending a single line of JavaScript scriplet into the Moodle application without the need for a website programmer to modify existing application code. Once the web page is loaded, the JavaScript tool is invoked registering all event handlers which are called for all events of interest triggered by the user when interacting with the interface. The events include typing, cursor movement and mouse clicking. Initially, the script transmits to the server side the initial state describing the *user* as well as the *website* information. Data related to the user includes the screen size, browser name, user default language and other details describing the session meanwhile information describing the website concerns primarily the page size and referral address.

3.3.3. Metrics estimation

Despite longstanding research in data extraction and mining, there is a dearth of automated methods for usability evaluation based on user interaction traces (Alspaugh, Ganapathi, Hearst, & Katz, 2014; Carta, Paternò, & Santana, 2011; Ivory & Hearst,

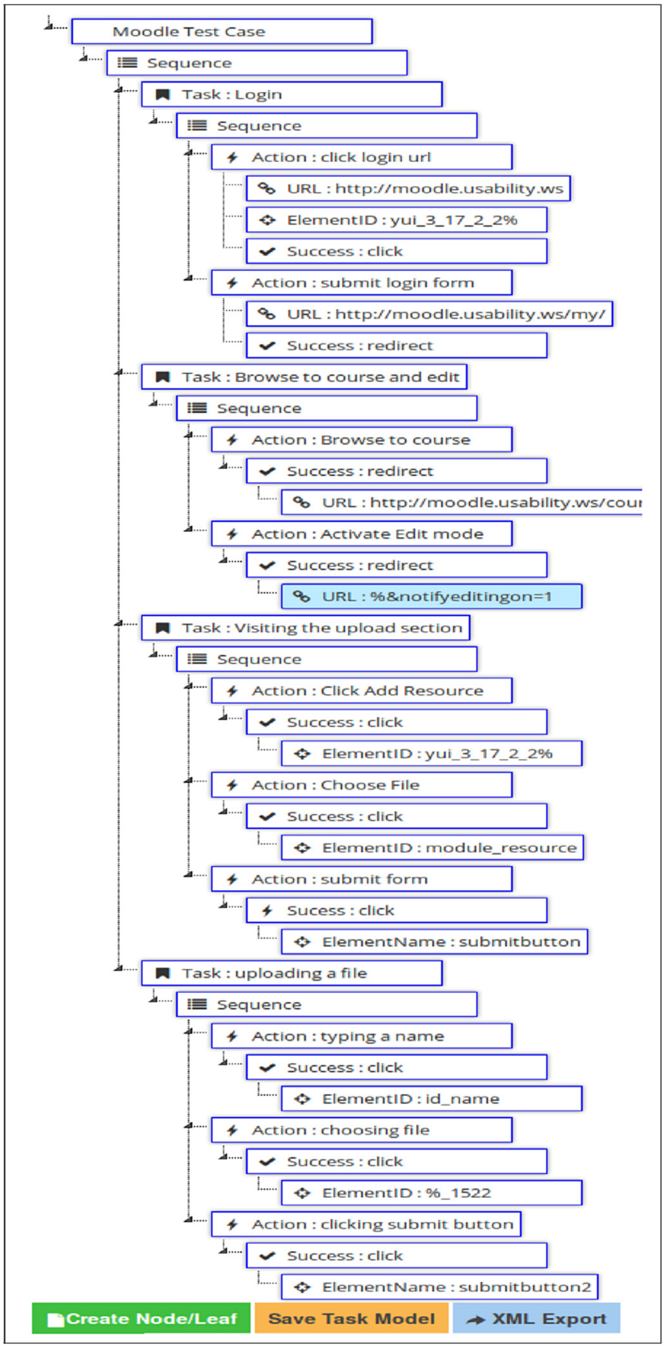

Fig. 2. Tree-based task descriptor for Moodle.

2001). As opposed to merely conducting analysis based on pre-defined anticipated usage measurements (Geng & Tian, 2015), problem identification for usability is driven via examining the deviation of derived metrics taken from the user logs through performing intra-class analysis across different dimensions to observe inconsistencies among diverse users. This is done together with matching the users data against measurements taken from experts considered as the ideal case to infer any heterogeneity between the designer and user cognitive models. Consequently, two types of logs are recorded for the *User* and *Expert* in order to infer usability issues for a given system. The tree-based task descriptor is thereafter being used to compare the constructed navigation trails which include all user interactivity actions such as recorded mouse clicks, cursor movement and inter-page navigations. The matching process is based on well-defined and conventional metrics that reflect better usability and assist with exploring the root causes of usability problems. The metrics are chosen on the basis that they can be quantified automatically without the cooperation of the participants purely based on the recorded client-side activity logs and the predefined task descriptor. The usability metrics considered in this research study include:

- *Cursor Distance (D)*: This is to measure the efforts undertaken by the user reflected through the use of hand to move the cursor on the screen. In practice, longer distances of the mouse cursor are a clear indication of poor usability and lower satisfaction level. The distance $D(t)$ for a specific task t is approximately estimated as the summation of the Euclidean distances between two consecutive log points for all N_t data activities enclosed for the given task as shown in the following Equation:

$$D(t) = \sum_{i=1}^{N_t - 1} \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2} \quad (1)$$

- *Completion rate (C)*: is also called the success rate which is considered one of the most fundamental usability metrics for a user to successfully achieve a goal. The completion rate $C(t)$ is typically measured as a binary percentage value for success (coded as 100%) or failure (coded as 0%) for primitive actions where it is impossible or impractical to define criteria for partial task success, but for simplicity reasons, binary values are considered for primitive actions. The completion of an action is inferred through the same way as the task duration via a sequential search for the *Success* condition. For an intricate task consisting of other structures and subtasks, the completion rate is estimated as the percentage of accumulated completion rates for sub tasks in a hierarchical fashion as expressed in the following Equation such that $||t||$ denotes the number of direct subtasks under the task t:

$$C(t) = \frac{\sum_{sub \in t} C(sub)}{||t||} \quad (2)$$

- *Task Duration (T)*: is defined as the total time taken to achieve a particular task by a user or usability expert. This metric is usually used to measure the efficiency rate as the temporal analysis can indicate valuable insight information for the HCI practitioner in order to evaluate the design of the web application such as average time spent on specific task by various users compared to other tasks. Based on the task descriptor, task duration $T(t)$ is approximated in seconds as the difference between of time-stamps for the start and last data logs for the task. The starting timestamps is usually deduced as the final timestamps for the previous task. The ending time is estimated through a sequential search within the user activity trail for the *Success* condition corresponding to the defined task.
- *Mouse Clicks*: In the same way as the cursor distance, this metric is described to measure the efforts taken by the end user when clicking or tapping their input pointing device as a mouse or touching pad. For simplicity, a single type of click is considered referring to both right or left meanwhile a double click is counted as two consecutive clicks.

3.4. System usability scale

In order to quantify the satisfaction of the university participants for using the e-learning platform, the System Usability Scale (SUS) (Brooke, 1996) is considered at the end of the experiment session. The SUS is one of the most popular methods in the literature which is devised mainly to evaluate the usability for web applications. Its popularity is gained among the HCI community mainly due to its desirable psychometrics including high reliability and validity (Bangor, Kortum, & Miller, 2008; Brooke, 1996; Lewis & Sauro, 2009). The SUS questionnaire is composed from ten questions with a mix of positive and negative items as shown in Appendix A. For each question, the respondent rates the magnitude of their agreement using a 5-point Likert scale with statements going from strongly disagree (1) to strongly agree (5). In order to compute the overall SUS score, the score contribution for each odd question which is positively worded the scale minus 1. For the even items, the score contribution is 5 minus the scale position. Therefore, each contribution ranges from 0 to 4. The SUS is the sum of all score contributions for the 10 items multiplied by 2.5 as shown in Equation (3) where U_i refers to the rating of the i^{th} item. The SUS scores ranges between 0 and 100 in 2.5-point increment where higher values reflect higher satisfaction from the user.

$$SUS = 2.5 \times \left[\sum_{n=1}^{5} (U_{2n-1} - 1) + (5 - U_{2n}) \right] \quad (3)$$

4. Analysis and results

During the experiment session, participants have been instructed to login to their online account for Moodle using the login credentials created for them in advance. Afterwards, they need to upload a given *pdf* file as a teaching resource into their assigned module. The university lecturers are not given any assistance or instructions of how to use the Moodle platform apart from giving them the login details and the web address as a special link customized to every user in order to detect and infer the user session. At the end of the evaluation session, users are asked to fill in the SUS questionnaire asking for their opinion about the ease of use when performing the different tasks on the e-learning platform. The subjective assessment of usability from questionnaires is based on the likert scale with values from 1 to 5 with 10 questions relating to the satisfaction, efficiency and effectiveness. The collected data is undoubtedly rich and complex to analyse due to the variety of methods and metrics. The study results are discussed and grouped into two categories with special reference to the strength and weakness of each approach.

4.1. Usability metrics evaluation

To inspect the usability of the e-learning Moodle platform, usage metrics are computed automatically based on the recorded traces of all participants. The metrics include number of clicks, task duration, cursor distance and completion rate. This is computed individually for every higher level task for every user trace through the prescribed task descriptor. In this evaluation experiment, we considered recording the data logs from three experts who are already familiar with the e-learning application. Table 2 shows a summary for the computed usability metrics and statistical values from all the data computed for both end users and experts. The mean value and standard deviation (SD) are provided across all dimensions for better analysis of usability. Surprisingly, from the list of 50 participants, a low completion rate of 10% for Task 4 is reported indicating that 5 lecturers only who managed to complete successfully the experiments meanwhile most users struggled to use the Moodle platform in order to upload a course lecture for their students. This reflects poor usability for the e-learning interface during the phases where lecturers could not complete the task.

Fig. 3 shows the summative results obtained based on the derived metrics for the four tasks. The user data is estimated as the mean of measurements derived automatically for the three shown dimensions: Task duration, cursor distance and mouse clicks. The error bars in the plot on the users data correspond to the standard deviations of the measurements. It is observed that there is always a considerable gap between the expert and users logged data with the expert having always lower values compared to the average user. For the case of task 2, there is a high variance among users in terms of time in addition to the fact that there is a remarkable difference between the expert and users which is the same for task 4. This can be an indicative to either to a usability deficiency of the designed interface at this phase of the application that needs to be addressed, or it can be related to the nature of the application and intellectual capacities of end users who may require more time to think for using the application. For the quantification of mouse clicks, the number of clicks is not consistent between the two parties for most of the cases with a remarkable high standard deviation among end users. In the same way, the cursor distance metric shows considerable differences between end users and expert data during all tasks with large variation between the two types of users which is explained by the complexity of the interface for the e-learning platform. Meanwhile, the obtained results presented in Table 3 show that there exists strong correlations between the discussed usage-based metrics across the performed tasks for all participants using the estimated Correlation Person r values.

4.2. SUS analysis

The system usability scale scores are computed for all lecturers who completed the 10-item questionnaires after conducting the experiment with the e-learning system. Table 4 shows the SUS scores placed against the estimated usage metrics for all users grouped by gender, different age groups and academic qualifications. The average usability score based on the subjective evaluation is reported to reach the value of 69.3% with a standard deviation of 14.81%. The Cronbach alpha α which refers to the reliability of measurements is estimated as 0.458 for all scores. This is an indicative for the strong reliability for questionnaire instrument used in the evaluation. In terms of usage metrics, the lecturers are observed to perform poorly struggling to the use e-learning platform compared to the expert measurements. Surprisingly, most users reported higher SUS scores expressing their satisfaction. Female lecturers showed greater self-content with the e-learning system against the male counterparts with minor differences in terms of usage metrics.

For analyzing the obtained results for the different age groups, younger users for the age group (25–29) are observed to have the highest completion rates whilst they expressed better satisfaction and acceptance for using the Moodle platform with a high SUS average score of 74.67% and a standard deviation of 9.95. Older people for the age group (50–65) reported similarly higher SUS score of 69.57% though the metrics obtained from their traces indicate poor performance during their interactivity. In fact, none of the participants from this age group have completed fully the experiment successfully. The reported results is consistent with the research findings confirmed by Bringula (2013). For the academic qualification, the results shows that university lecturers with the rank of full professor have expressed higher satisfaction

Table 2
Summative measures for usability metrics per task.

	User completion rate (C) %	Duration			Distance			Clicks		
		Expert	User		Expert	User		Expert	User	
		Mean	Mean	SD	Mean	Mean	SD	Mean	Mean	SD
Task 1	93.0	25	234	380	6.9	15.8	25.6	5.0	9.1	9.7
Task 2	51.0	10	512	633	2.3	40.0	47.7	2.0	26.1	33.7
Task 3	25.33	11	150	170	1.7	12.9	16.9	5.0	13.4	15.1
Task 4	10.0	22	138	68.2	2.8	10.5	39.8	2.0	11.2	7.1

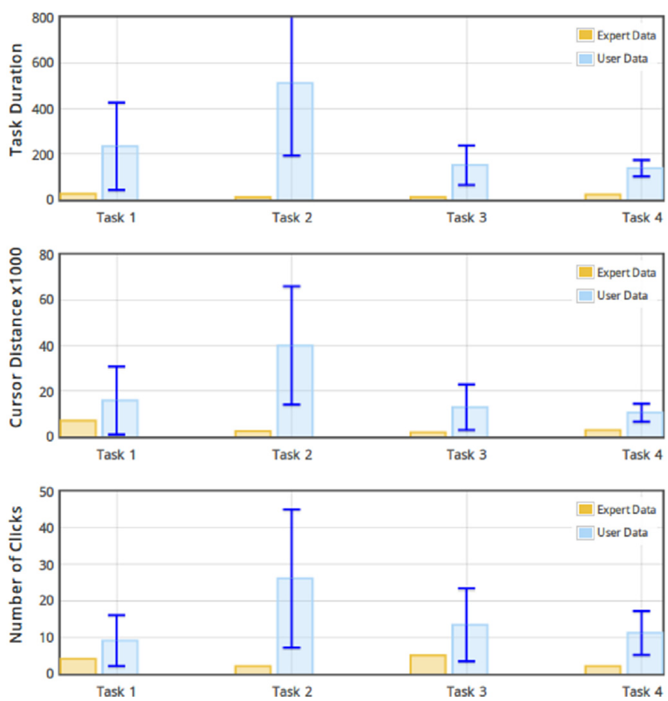

Fig. 3. Usage metrics per task via client-side traces.

Table 3
Pearson Cross-Correlations for SUS & usage metrics.

	1	2	3	4	5
1. Complet. Rate	–				
2. Task Duration	−0.276	–			
3. Cursor Dist.	−0.46	0.863	–		
4. Mouse Clicks	−0.472	0.762	0.753	–	
5. SUS Score	0.257	0.036	−0.024	−0.227	–

rates via SUS analysis whilst all of them failed to perform the tasks assigned to them. Inversely, there are consistencies between the expressed satisfaction level and interaction performance achieved by assistant lecturers. Fig. 4 shows the relationship between the completion rates and the expressed SUS scores by university lecturers with $F(1,98) = 32.44$ and p value of 0.00013 obtained using Anova analysis taking into account the two variables.

5. Discussions

This research study makes a number of unique contributions to the evaluation of e-learning systems and exploring the factors that can affect the satisfaction level and interactivity performance for university lecturers for using educational technologies. Primarily, the results obtained via the conducted experiments confirm that the System Usability Scale score is not a sufficient measure to express the true acceptance and satisfaction level of lecturers for using the e-learning systems. The evaluation must be fulfilled in tandem with analyzing the usage metrics derived from interactivity traces in a non-intrusive fashion. This would help to cluster different users and even conduct deep analysis of the reported usability scores based on the actual performance of participants. Consequently, the reported satisfaction scores reported by means of questionnaires administered to a set of users, can be potentially have different interpretation by the user in expressing their acceptance level. In other words, are the lecturers satisfied because of the ease of use for the e-learning platform or because of experiencing a new technological product that they felt happy about it regardless of the expected results. The same argument has been confirmed in a recent study about the relationship between user ratings versus their expectations (Michalco, Simonsen, & Hornbæk, 2015). Based on the attained satisfaction scores in comparison to the usage data, the recommendation for assessing the satisfaction level based on the ease of use would be to let the users aware of their completion success rates or instead start the questionnaire with further questions asking the users if they have completed the task successfully.

The results have shown that the different usage-metrics including: Task duration, number of clicks, completion rate and cursor distance play equivalently the same role in expressing and analyzing the usability degree of user interactivity. For other factors related to the participants themselves, younger users have shown greater motivation and skills to use technological products meanwhile older users have struggled poorly to use the e-learning platform. This was based on the computed usage metrics regardless of the reported ratings. This is in alignment with a number of recent studies which arrived to the same conclusions (Bringula, 2013; Wagner, Hassanein, & Head, 2014) arguing that the age factor has a pronounced impact on the performance of users. In the same way, lecturers with the highest academic qualifications have reported declining performance with low completion rates. This is intuitively due to the proportional relationship between the age and the academic qualification. Conversely to other studies (Mentes & Turan, 2012; Page, Robson, & Uncles, 2012) that argued that gender is a factor that impacts the performance of users, the results obtained in this study assert that both genders have almost the same usage metrics with marginal differences with the exception that female lecturers have expressed greater self-content with the e-learning system.

As this study is not to explore the functionalities provided by the open e-learning platform Moodle, but rather to assess the usability of its interface. The usage-based metrics have shown that the university lecturers have struggled to interact with the e-learning system when facing web pages with many graphical widgets and options. This suggests the partial poor usability of the user interface

Table 4
Usability Metrics & SUS scores per Gender, Age Groups and Academic Qualifications.

		Usage-based metrics							SUS scores	
		Completion rate	Duration		Distance		Clicks			
			Mean	SD	Mean	SD	Mean	SD	Mean	SD
Experts		100	35.7	5.13	9.44	1.17	7	0	–	–
All Lecturers		45.33	385	254	44.28	35.89	22	12.21	69.30	14.81
Gender	Male	50.15	390	266	36.23	24.77	23	11.2	68.70	14.62
	Female	39.67	378	255	36.23	24.77	16	27.21	70.00	15.65
Age Groups	25–29	42.19	463	315	34.71	45.4	26	21.5	74.67	9.95
	30–39	50.24	366	393	26.61	43.32	29	21.9	70.15	15.52
	40–49	42.08	214	194	17.64	7.38	13	8.96	59.75	21.42
	50–65	45.24	498	158	35.55	11.97	32	15.3	69.57	6.61
Acad. Qualif.	Professor	41.67	498	97	41.59	13.10	32	13.43	68.93	7.20
	Lecturer	50.78	271	244	24.83	13.32	16	8.43	67.03	21.78
	Assit. Lecturer	43.06	429	326	59.89	48.15	31	24.47	70.74	11.63

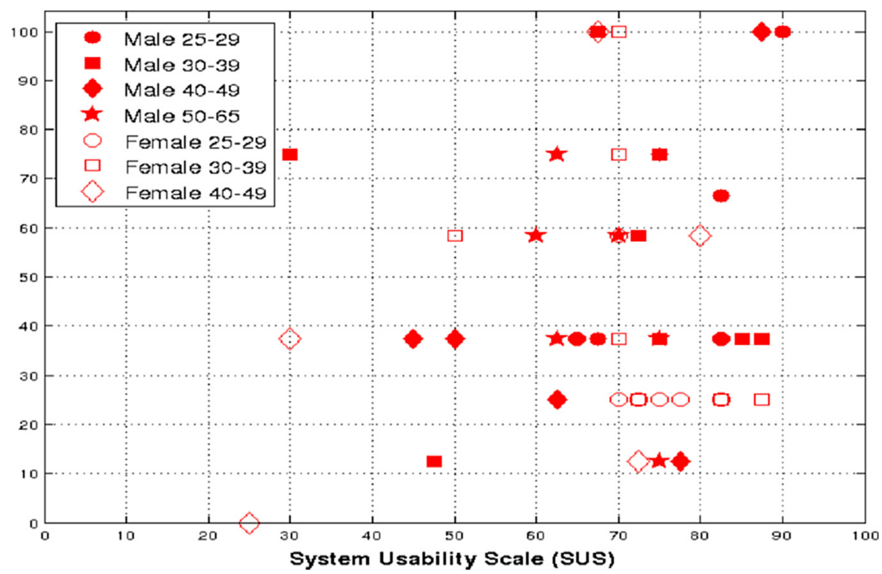

Fig. 4. Relationship between Completion Rate vs. SUS sores.

that must be improved during the phases where most users failed to complete the tasks. Meanwhile, minimal interfaces are proven to be better in terms of achieving goals with ease and consistency concluding the correlation between the complexity of the task and the number of mouse clicks, duration and cursor distance with respect to the number of elements and options contained within the interface. Furthermore, users have expressed their satisfaction to using the system in the future for online teaching whilst they have demanded explicitly more training and guidance of how to use the e-learning platform.

6. Conclusions

Positive user experience and better usability are of prime importance for educational-based learning systems playing vital role for the acceptance, satisfaction and efficiency of academic institutions. Empirical-based study is conducted to explore how lecturers interact with an e-learning environment system based on a predefined task model describing low-level interactivity activities for uploading teaching resources for students. To evaluate usability for a given website and infer further insights about the ease of interactivity, various metrics including cursor distance, mouse clicks and task duration are computed automatically. The experiments have revealed that the System Usability Scale is not adequately a standalone measure for expressing the true acceptance and satisfaction of lecturers for using the e-learning systems. The evaluation must be fulfilled together with analyzing the automated metrics derived from usage activities. For future avenues to explore within this research, the usability of other modules of the e-learning platform can be assessed such as assignments, quiz and forums. Furthermore, the memorability and learnability aspects with their relationship to the ease of use can be investigated through analyzing the usage logs of users throughout time.

Acknowledgment

The authors would like to thank all the university lecturers who were willing to take part in the usability evaluation project for e-learning systems.

Appendix A

System usability scale
1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

References

Alspaugh, S., Ganapathi, A., Hearst, M. A., & Katz, R. (2014). Better logging to improve interactive data analysis tools. In *KDD workshop on interactive data exploration and analytics*.

Asarbakhsh, M., & Sandars, J. (2013). E-learning: the essential usability perspective. *The Clinical Teacher, 10*, 47–50.

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human–Computer Interaction, 24*, 574–594.

Bringula, R. P. (2013). Influence of faculty-and web portal design-related factors on web portal usability: a hierarchical regression analysis. *Computers & Education, 68*, 187–198.

Brooke, J. (1996). Sus-a quick and dirty usability scale. *Usability Evaluation in Industry, 189*, 4–7.

Capece, G., & Campisi, D. (2013). User satisfaction affecting the acceptance of an e-learning platform as a mean for the development of the human capital. *Behaviour & Information Technology, 32*, 335–343.

Carta, T., Paternò, F., & De Santana, V. F. (2011). Web usability probe: a tool for supporting remote usability evaluation of web sites. In *Human-computer interaction–INTERACT 2011* (pp. 349–357). Springer.

Cassino, R., & Tucci, M. (2011). Developing usable web interfaces with the aid of automatic verification of their formal specification. *Journal of Visual Languages & Computing, 22*, 140–149.

Cassino, R., Tucci, M., Vitiello, G., & Francese, R. (2015). Empirical validation of an automatic usability evaluation method. *Journal of Visual Languages & Computing, 28*, 1–22.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management Science, 35*, 982–1003.

Dominici, G., & Palumbo, F. (2013). How to build an e-learning product: factors for student/customer satisfaction. *Business Horizons, 56*, 87–96.

Escobar-Rodriguez, T., & Monge-Lozano, P. (2012). The acceptance of moodle technology by business administration students. *Computers & Education, 58*, 1085–1093.

Fernandez, A., Insfran, E., & Abrahão, S. (2011). Usability evaluation methods for the web: a systematic mapping study. *Information and Software Technology, 53*, 789–817.

Garrison, D. R. (2011). *E-learning in the 21st century: A framework for research and practice*. Taylor & Francis.

Geng, R., & Tian, J. (2015). Improving web navigation usability by comparing actual and anticipated usage. *Human-Machine Systems, IEEE Transactions on, 45*, 84–94.

Guerrero Garcia, J., Vanderdonckt, J., & Gonzalez Calleros, J. M. (2008). Flowixml: a step towards designing workflow management systems. *International Journal of Web Engineering and Technology, 4*, 163–182.

Guri-Rosenblit, S. (2006). Eight paradoxes in the implementation process of e-learning in higher education. *Distances et Savoirs, 4*, 155–179.

Hornbæk, K. (2006). Current practice in measuring usability: challenges to usability studies and research. *International Journal of Human-Computer Studies, 64*, 79–102.

Ivory, M. Y., & Hearst, M. A. (2001). The state of the art in automating usability evaluation of user interfaces. *ACM Computing Surveys, 33*, 470–516.

Laurillard, D., Oliver, M., Wasson, B., & Hoppe, U. (2009). Implementing technology-enhanced learning. In *Technology-enhanced learning* (pp. 289–306). Springer.

Lee, S. M., Kim, Y. R., & Lee, J. (1995). An empirical study of the relationships among end-user information systems acceptance, training, and effectiveness. *Journal of Management Information Systems*, 189–202.

Lewis, J. R., & Sauro, J. (2009). The factor structure of the system usability scale. In *Human centered design* (pp. 94–103). Springer.

Li, Y., Duan, Y., Fu, Z., & Alford, P. (2012). An empirical study on behavioural intention to reuse e-learning systems in rural China. *British Journal of Educational Technology, 43*, 933–948.

Marco, F. A., Penichet, V. M. R., & Gallud, J. A. (2013). Collaborative e-learning through drag & share in synchronous shared workspaces. *J. UCS, 19*, 894–911.

Mentes, S. A., & Turan, A. H. (2012). Assessing the usability of university websites: an empirical study on namik kemal university. *Turkish Online Journal of Educational Technology-TOJET, 11*, 61–69.

Michalco, J., Simonsen, J. G., & Hornbæk, K. (2015). An exploration of the relation between expectations and user experience. *International Journal of Human-Computer Interaction, 31*, 603–617.

Navimipour, N. J., & Zareie, B. (2015). A model for assessing the impact of e-learning systems on employees satisfaction. *Computers in Human Behavior, 53*, 475–485.

Nichols, M. (2008). Institutional perspectives: the challenges of e-learning diffusion. *British Journal of Educational Technology, 39*, 598–609.

Orfanou, K., Tselios, N., & Katsanos, C. (2015). Perceived usability evaluation of learning management systems: empirical evaluation of the system usability scale. *The International Review of Research in Open and Distributed Learning, 16*.

Page, K. L., Robson, M. J., & Uncles, M. D. (2012). Perceptions of web knowledge and usability: when sex and experience matter. *International Journal of Human-Computer Studies, 70*, 907–919.

Persico, D., Manca, S., & Pozzi, F. (2014). Adapting the technology acceptance model to evaluate the innovative potential of e-learning systems. *Computers in Human Behavior, 30*, 614–622.

Phillips, R., McNaught, C., & Kennedy, G. (2012). *Evaluating e-learning: Guiding research and practice*. Routledge.

Quade, M., Lehmann, G., Engelbrecht, K.-P., Roscher, D., & Albayrak, S. (2013). Automated usability evaluation of model-based adaptive user interfaces for users with special and specific needs by simulating user interaction. In *User modeling and adaptation for daily routines* (pp. 219–247). Springer.

Renaut, C., Batier, C., Flory, L., & Heyde, M. (2006). *Improving web site usability for a better e-learning experience. Current developments in technology–assisted education* (pp. 891–895). Badajoz, Spain: FORMATEX.

de Santana, V. F., & Baranauskas, M. C. C. (2015). Welfit: a remote evaluation tool for identifying web usage patterns through client-side logging. *International Journal of Human-Computer Studies, 76*, 40–49.

Simões, A. P., & de Moraes, A. (2012). The ergonomic evaluation of a virtual learning environment usability. *Work-Journal of Prevention Assessment and Rehabilitation, 41*, 1140.

Stoffregen, J., Pawlowski, J. M., & Pirkkalainen, H. (2015). A barrier framework for open e-learning in public administrations. *Computers in Human Behavior, 51*(Part B), 674–684.

Tsai, P. C.-F., Yen, Y.-F., Huang, L.-C., & Huang, C. (2007). A study on motivating employees learning commitment in the post-downsizing era: job satisfaction perspective. *Journal of World Business, 42*, 157–169.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: four longitudinal field studies. *Management Science, 46*, 186–204.

Violante, M. G., & Vezzetti, E. (2015). Virtual interactive e-learning application: an evaluation of the student satisfaction. *Computer Applications in Engineering Education, 23*, 72–91.

Wagner, N., Hassanein, K., & Head, M. (2014). The impact of age on website usability. *Computers in Human Behavior, 37*, 270–282.

Yeh, Y.-c., & Lin, C. F. (2015). Aptitude-treatment interactions during creativity training in e-learning: how meaning-making, self-regulation, and knowledge management influence creativity. *Journal of Educational Technology & Society, 18*, 119–131.