Usability Testing of MOOC: Identifying User Interface Problems

Olga Korableva\textsuperscript{1,2}, Thomas Durand\textsuperscript{3,5}, Olga Kalimullina\textsuperscript{4} and Irina Stepanova\textsuperscript{5,6}

\textsuperscript{1}St. Petersburg State University, S-Petersburg, Russian Federation
\textsuperscript{2}Institute of Regional Economic Studies of Russian Academy of Science, Leading Researcher, S-Petersburg, Russian Federation
\textsuperscript{3}Conservatoire National des Arts et Métiers, Paris, France
\textsuperscript{4}The Bonch-Bruevich St. Petersburg State University of Telecommunications, St-Petersburg, Russian Federation
\textsuperscript{5}ITMO University, St-Petersburg, Russian Federation

Abstract: In the modern world, more and more information systems are actively used in the educational process. Examples of such systems are platforms for hosting remote MOOC (massive online open course). However, multiple MOOCs are perceived differently by users and have various levels of completion. It was proved that the existing problems of such systems are related to the usability of their user interface. A number of techniques are used to investigate user satisfaction with the interface. Most of them evaluate, first of all, the user satisfaction index after the course completion or at the stage of prototype creation and testing. The authors of the article carried out a research of existing approaches and proposed their own methodology for the evaluation of user satisfaction with the interface design on the basis of questionnaires UMUX-Lite, SUS, Testbirds Company approach and the ISO standards. The study allowed identifying gaps in the design of each of the analyzed platforms and its perception by users.

1 INTRODUCTION

MOOC is one of the interactive learning tools offering online courses that allow students to access learning resources anytime and anywhere. Therefore, the MOOC is now becoming increasingly popular all over the world. Gay and Go researchers (2018) based on a survey of Chinese students identified the relationship of targets and tools that characterize the value of MOOCs for the educational process. According to the results, the effectiveness of the application in training, experience and usability of MOOC are identified as the main goals to maximize the importance of the platform in education.

Of course, the opportunity to receive education remotely and at your own pace has many advantages. However, the MOOC has a number of drawbacks, some of which are the result of problems with the user interface of online courses. So, according to some reports, the percentage of completion of the online course ranges from 7-20\% (Anderson et al., 2014), partly due to user dissatisfaction with the interface design. Thus, the problem of improving the usability of the MOOC user interface is of great practical importance (Sethi, 2017). High-quality design and minimization of “defects” of the platform is a way to solve this problem. An important part of the system evaluation is the user experience. In the process of interaction with the system, it is possible to reveal awareness, emotions, physiological and psychological behavior of a user (Kuhlthau, 1991; ISO 9241-210). However, most studies of usability issues apply technologies to collect data during direct interaction with the MOOC platform (Hu, 2019; Iniesta and Rodrigo, 2018; Liu et al., 2018; Cheng et al., 2018; Gao et al., 2018; Maloshonok and Terentev, 2016) or at the stage of its development (Morales and Benedí, 2017). However, it is not always possible to

\begin{itemize}
  \item \url{https://orcid.org/0000-0002-2699-8396}
  \item \url{https://orcid.org/0000-0002-5122-0746}
  \item \url{https://orcid.org/0000-0002-7782-6148}
  \item \url{https://orcid.org/0000-0002-8552-246X}
\end{itemize}
evaluate a user's response in direct contact with the environment under study. Thus, the issue of creating an adapted methodology to assess the user interface satisfaction without direct contact with the platform becomes relevant, because it can significantly increase the potential number of participants and ensure greater representativeness of the data.

2 THEORETICAL BACKGROUND

Usability research is becoming increasingly relevant both in various areas of life and in various scientific fields. There are numerous studies in the field of IT-technologies, where the results of usability assessment are indisputable. Neglecting such an assessment at the design stage adversely affects the usability of any software. Various mobile applications and platforms are evaluated (Faurholt-Jepsen et al., 2019), the usability of software systems for social data analysis and for extracting useful knowledge from social networks user data is assessed (Wang et al., 2019). Many studies are devoted to the issues of online trading usability, such as, for example, virtual fitting rooms applications (Jo and Kim, 2019) explores the potential usability of the automated structure of a public data set for machine learning, soft computing and cybersecurity (Martín et al., 2019), studies the relationship between user interaction and digital libraries evaluation (Li and Liu, 2019). There are also studies on the development of software management tools used to enable the evaluation of usability activities in a flexible environment (Deraman and Salman, 2019), and other studies.

It is necessary to mention that usability assessment has a great social function. For example, the principles of designing the user interface for mobile applications for the convenience of seniors are being studied (Wildenbos et al., 2019).

On the whole, usability assessment is relevant for a wide range of areas: from evaluation of the robots’ usability as a part of smart home (Wilson et al., 2019) to identification of critical quality dimensions for continuance intention in mHealth services (Kim et al., 2019). There is also a large number of usability studies in the field of education: the study of virtual reality (VR) technologies’ usability within the educational process (Makransky and Petersen, 2019), an analysis of online exams usability problems (Ulha and Ali Babar, 2019), and other are of particular interests (Álvarez-Xochihua et al., 2017).

2.1 Usability Requirements and Quality Standards

The regulation and evaluation of the quality indicators of the platform's design are based on ISO standards. For example, ISO / IEC 9126-3 regulates the internal indicators of platform usability; ISO / IEC 9126-2 – defines external quality indicators; ISO / IEC 9126-4 – regulates the platform usage quality indicators.

All applied standards are largely based on determining the quality of software or information product based on the ability of a particular product to help specific users achieve certain goals efficiently, quickly and safely. Thus, the required interface design implies ease of use, functionality, efficiency, and reliability at the same time.

The main usability requirements include:

1. Performance, which means that the task is performed by users with an accuracy of at least 95% in less than 10 minutes.
2. User satisfaction. The following are methods for assessing the usability/interface quality: the MUSiC performance evaluation method, the SUMI questionnaire, the usage context assessment, the actual context of assessment and other methods.

2.2 Methods and Principles of User Interface Design

In order to meet the requirements of usability, when designing a UX interface it is advisable to use the methods and principles used in the implementation of human-computer interactions (human-computer interaction – HCI).

1) Anthropomorphic Approach – assumes the development of a user interface as a system with qualities similar to human ones. Communication of the system with the user is built like a person to person interaction.
2) Cognitive Approach – considers the possibilities of the human brain and sensory perception of a person in order to develop a user-friendly interface.
3) Empirical Approach – is used to study and compare several concepts of interface design. Users evaluate specific elements of one complex concept in terms of usability.
4) The Predictive Approach is Associated with the GOMS (goals, objects, methods, and selection rules) acronym and means methods of studying the
individual components of user experience in terms of time it takes the user to achieve the goal. Goals reveal the user’s ultimate objective on the website.

It should be noted that, since the user’s goals can be completely different, then, in case of online courses, the UX designer should have a common understanding of the educational platform listener’s behavior and also a predetermined pattern of user behavior on a specific information page. Anderson et al. (2014) conducted an analysis of student behavior during the Coursera platform courses. The authors of the article using cluster analysis of data on the interaction of users with the online platform found that the online educational platform user behavior can be described by several common patterns that include certain types of actions performed by the listener with a certain frequency on the online platform.

At the same time, other researchers (Rodrigues et al., 2016a), who analyzed the Open Edu platform, conducted a similar cluster analysis of the total user actions on the platform. They identified only three categories of users: involved, periodically involved and not involved. And the latter category included the largest number of students.

2.2.1 Interface Design Research Methods

Vermeeren et al. (2010) evaluated 96 methods for interface design (UX design) research. Most of the studied methods, according to the authors of the article, can be used for the last stages of product development (prototype creation and prototype testing). For example, only a few of the 96 user interface analyzing methods are suitable for platform auditing, since they allow an assessment of an already completed website/platform with the involvement of third-party users. However, they cannot be applied at the stages of creating and testing an interface prototype.

Direct interaction between the user and the computer is necessary for the implementation of most techniques. For example, Foraker Labs applies the Heuristic Evaluation approach in order to assess the interface design. The heuristic evaluation (usability audit) is an interface evaluation by one or more experts. Only SUS and UMUX-Lite techniques can be used in order to evaluate user interface satisfaction, being the most commonly used methods to study user interface convenience and ways to obtain user experience and user satisfaction data. What makes these techniques unique is that they allow creating a questionnaire with graphic elements, notably print screens selected in a particular way, which in turn allow evaluating the platform without directly interacting with it.

3 METHODOLOGY AND PARTICIPANTS

3.1 Methodology

Two closely related courses on the Coursera and Open Education platforms were considered as part of the study.

The following assessment options were included when developing a methodology for studying the user interface satisfaction of online education platforms:

1. availability of actions that a user can take with an object (designation of actions that can be performed with labels, buttons, icons, scroll bars, etc.),
2. assessments of metaphorical design — an effective way to transfer abstract information that allows users to understand the meaning of actions they can perform with an object (example: desktop, recycle bin on personal computers; metaphors allow users to quickly learn how to use the system).
3. consideration of information processing models and the cognitive load of a person. It was taken into account that at first, a person perceives any information through the senses using his/her sensory system (hearing, sight, smell, touch), then transmits this information into short-term memory and holds it there in a limited amount for 30 seconds. Then the data goes into long-term memory or is forgotten. After the information “leaves” for the long-term memory, it can be called or recognized through similar objects from the outside. It is also important to consider the level of user attention of the user when designing an interface. As a rule, the user can focus only on one task at a specific point in time. Too many response options can make the user feel uncomfortable and may even cause the desire to leave the resource without achieving the goal.

User’s goals were also analyzed during the development of the methodology. Based on research by Anderson et al. 2014; Rodrigues et al., 2016b, Rieber, 2017, who analyzed the behavior of students at the Coursera and Openedu online educational platforms courses, developed the general structure of user behavior pattern:

- Involved: “Universals”. 
Periodically involved: “Spectators”, “Solvers”, “Collectors”.
Not involved: “Observers”.
Most likely, for people not involved in online education, the interface of the educational environment will not play an important role. For those who are periodically involved, only certain structural elements that correspond to their platform behavior pattern will be significant (page design while taking tests and examination tasks for solvers, page layout for viewing lectures for spectators, etc.).

The resulting total classification of user behavior patterns was used in the study of user interaction with the platform interface at the stage of processing the results of this study.

**Compiling a questionnaire.** The questionnaire integrated the questions for system usability scale (SUS) and Usability Metric for User Experience – Lite (UMUX-Lite) methodologies and Testbirds Company approach.

Initially, the SUS methodology is a survey of respondents using a questionnaire with an opportunity to arrange the most acceptable answer on a scale from 1 to 5 (from “strongly disagree” to “absolutely agree”). In total, there are 10 questions in the questionnaire. The questions based on the SUS methodology, that concern such spheres as the desire to use the system frequently, the system’s perception as complex / simple, the need to involve a technician in order to use the system, presence of inconsistency in the system, the speed of mastering the system, perception of system as cumbersome, confidence in using the system, the need to study additional material to facilitate the interaction with the system were chosen to be included in the questionnaire.

After the survey, all the answers to questions based on the SUS method were evaluated according to the following rules:

1) Answers to questions numbered 1,3,5,7, 9 receive a number equal to the scale number minus one.
2) Answers to questions numbered 2,4,6,8, 10 receive a number equal to 5 minus the number of the scale.
3) After revaluation, all values are summed up and multiplied by a 2.5 coefficient.

Also, the questionnaire included questions based on the UMUX-Lite methodology – a methodology for assessing user interface satisfaction, which was developed first based on the UMUX assessment methodology, and even earlier – based on the SUS. The questions about matching the system capabilities with the requirements of the user, perception of using the system (frustration/delight), ease of the system usage and time spent on understanding the operation of the system were chosen to be included in the questionnaire. Questions based on the UMUX-Lite methodology have a scale from “strongly disagree” to “absolutely agree”, but ranges from 1 to 7. The first two answers are evaluated as follows: the answer value minus one, then all indicators are summed up, divided by twelve and multiplied by 100 (Lewis etc., 2013).

The general formula by which the regression dependence is calculated in UMUX-Lite:

\[ \text{UMUX-LITE} = 0.65\times([\text{Item 1 score}] + [\text{Item 2 score}] - 2)/100/12) + 22.9. \]

For a single calculation case, the coefficients of 22.9 and 0.65 are not taken into consideration.

Also, some of the questions were formulated according to the Testbirds Company’s methodology. These questions revealed four types of defects that may be present on the platform:

- **Functional defect** – refers to certain functions of the test object. (example: you cannot click a button, you cannot open a drop-down list).
- **Display defect** – a defect of incorrect displaying of information/media files/widgets (example: an unreadable character, duplication of descriptions in columns, some interface elements are displayed with distortions).
- **Performance defect** – freeze/crash/malfunction of test objects occur (the application is slow on a mobile device certain objects are loaded poorly).
- **Spelling errors** – any errors that are not consistent with the rules of the language.

Defects are assessed on a single scale: critical – blocking error, which causes the application to be inoperable; High – an incorrectly working key object of the system, the incorrect operation of which results in the non-working state of a certain part of the system, without the possibility of solving the problem Medium – a significant error when part of the main business logic does not work correctly. The error is not critical and allows working with the function under test using other methods. Low – a minor error that does not violate the business logic of the tested part of the application, an obvious user interface problem, or an error not related to business logic.

Thus, based on the analysis of existing approaches to assessing user interface satisfaction, a methodology for assessing user behavior when
signing up for the course and user behavior during the course was developed. Screenshots were created for each stage of user interaction with the platform. A mental model diagram was formed on the basis of the received packages of images, UMUX-Lite and SUS questionnaires, the Testbirds approach, and ISO. The mental model diagram was useful for identifying gaps in the design when the system did not fully satisfy users’ needs and became the basis for building the final questionnaire. The developed questionnaire for collecting data on assessing users interface design satisfaction contains 40 questions, 4 of which are common, the remaining 36 concern two of the studied platforms equally (18 questions about the Coursera platform and 18 questions about the Open Education platform).

The developed method makes it possible to assess the information environment by several probable criteria:

• Correlation of the system with the real world of the user (use of words, terms, and modules corresponding to the level of development, education, and behavioral features of the target audience).

• Presence of user freedom and control — the ability of the user to independently control his actions in the system. For example, the ability to cancel or repeat actions, as well as to log out at any time.

• Consistency and unity — all interface elements: icons, terminology, error messages and so on should be uniform and consistent throughout the interface. The use of generally accepted icons allows users to quickly master the system during their initial acquaintance.

• User error prevention is a concept in which the system warns the user in case of irreversible actions. It is also important to give users the ability to undo such actions by storing data in the database as well as to warn them about entering incorrect data or completing forms incorrectly.

• Automatic data loading, cognitive load reduction — the ability to help the user to perform actions easier. For example, by loading previously entered information or offering auto-input for some fields.

• Flexibility and efficiency of the system for experienced and new users. For example, you can enter hot keys to speed up user interaction with the system.

• Minimalistic design — the principle involves using the minimum number of background illustrations when displaying important material.

• Help and documentation — it is assessed how visually easy the user’s documentation is found (for example, it is useful to provide users with video tutorials on complex elements, etc.) (User experience modeling., N.d.).

3.2 Participants

The work had the interface attractiveness for users and the predictive behavior of users on the example of an analytical comparison of two online learning platforms, Coursera, and Open Education, were studied. The Open Education Platform is a Russian online educational platform created in December 2014 with the support of the Ministry of Education of Russia and 8 basic universities of the country. The platform provides an opportunity to create a personalized learning path, to form an electronic portfolio and to purchase paid courses. The Coursera, global educational platform, was founded in 2012, and as for the end of 2017, the number of registered users of this resource was 30 million.

In order to study the users’ perception of the online courses interface on both platforms, it was necessary to find out the target audience of the projects. The SimilarWeb analytics service was used for analyzing the target audience of those projects. Age and gender data of the Open Education website visitors are presented in Figure 1. Age and gender data of the Coursera website visitors is provided in Figure 2.

Figure 1: Age and gender data of the Open Education resource visitors.

Figure 2: Age and gender data of the Coursera resource visitors.

Based on the data received we can conclude that the main target audience of these two educational platforms consists of users from 18 to 35 years. Since the gender of the target audience of the two platforms is different, we will use an arbitrary number of users of one or the other gender
in our study. However, it was decided not to place an upper limit for the age category.

Participants of the experiment were respondents over 18 years old. A total of 60 people were interviewed: out of those, 58.3% are between the ages of 18 and 24, 31.7% are 25-35 years old and 10% are over 36 years old. There were 61.1% of women and 38.3% of men among the respondents. Most of the respondents (71.7%) used online educational platforms before taking the survey.

4 RESULTS

Based on the author's methodology, a mental model diagram was built (Fig. 3), according to which a questionnaire was formed and a survey was conducted. The developed mental model diagram made it possible to design a questionnaire reflecting the general structure of the user behavior pattern.

According to the assessment of the interface using adjectives that describe the perception of user interaction with the platform, the following descriptive characteristics were obtained – The comparative frequency of correlation of an adjective and the interface of the corresponding platform is presented in Fig. 4.

As a result of studying the satisfaction from online platforms interface, data were obtained in relative scales for the Open Edu and Coursera platforms. The data of the obtained satisfaction calculations by users of the two platforms are presented in Fig. 5. According to the literature (Borsci et al., 2015), the obtained indicators are classified as follows:

- Open Education – to interface with Grade F,
- Coursera – to interface with Grade D.

In this case, the scale starts from A + absolutely satisfactory to Grade F – absolutely unsatisfactory.

The UMUX and SUS methodology tests results do not have to be equal, but most often they are close to each other. In this case, the evaluation of these methods showed similar results (Fig. 5).
Therefore, we can conclude that users do not find any platform difficult, but Coursera was rated as less complicated when comparing two platforms. In general, both platforms appear in good shape: users highly appreciated both platforms in terms of simplicity and accessibility, amenity and creativity.

Also in the course of this study the user behavior pattern data was obtained. Thus, the overwhelming majority of survey participants were equally eager to participate in solving problems and watching video lectures, which means they are “universals”.

Recommendations for changing the user interface of the analyzed platforms were developed in the framework of the study. Thus, the Open Education platform is recommended to work on a simpler and intuitive organization of graphic and textual material, implementing modern templates that can increase the loyalty of the main target audience using modern fashionable design. For the Coursera platform, it is also important to revise the main interface using innovative solutions, which would allow raising the level of user loyalty in the “fashionable, innovative, modern” directions.

5 CONCLUSION

It was intended to analyze already existing user interfaces within the framework of this study, so the research methodology was developed first, then the study of user reaction to the interface of educational online platforms was carried out, and the type of behavioral factor was determined and certain recommendations were formulated for the development of designated educational platforms. Thus, the main objective of this scientific work was to develop a methodology for researching online education platforms without direct user contact with the platform and then test the developed methodology. Based on the analysis of existing methodologies, the authors formed a mental model diagram, and then a questionnaire for collecting data on assessing satisfaction from user interface design based on UMUX-Lite, SUS questionnaires, the Testbirds Company’s approach, and the ISO. The data obtained allowed identifying gaps in the design of Coursera and Open Education platforms.

The proposed research methodology was developed for those who need to conduct a third-party assessment of the online educational platform interface design with no access to the internal metrics of the resource and also without the need to involve respondents in the online course. The methodology simplifies the researchers’ task by providing respondents only with an assessment questionnaire, which contains all the necessary and significant criteria and conditions.

ACKNOWLEDGEMENTS

This research is supported by RFBR (grant 16-29-1296518).

REFERENCES


